GENETIC AND EPIDEMIOLOGICAL CHARACTERIZATION OF *STREPTOCOCCUS PNEUMONIAE* DISEASE DETERMINANTS.

Sara Melchiorre

Dottorato in Scienze Biotecnologiche – XXIV ciclo

Indirizzo Biotecnologie Molecolari ed Industriali

Università di Napoli Federico II

Dottorato in Scienze Biotecnologiche – XXIV ciclo

Indirizzo Biotecnologie Molecolari ed Industriali

Università di Napoli Federico II

GENETIC AND EPIDEMIOLOGICAL CHARACTERIZATION OF *STREPTOCOCCUS PNEUMONIAE* DISEASE DETERMINANTS.

Sara Melchiorre

Dottoranda: **Sara Melchiorre**

Relatore: **Prof. Ettore Benedetti**

Coordinatore: **Prof. Giovanni Sannia**

Table of contents

Study ethics

Results

Pneumococcal isolates and PI-1 prevalence

PI-1 and transmission

PI-1 and first pneumococcal carriage duration in infants

Discussion

References

Chapter 2

Title: "Point mutation in *wchA* of 7F capsular locus are responsible for non-typeability of two invasive *Streptococcus pneumoniae* isolates".

Introduction

Materials and Methods

Bacterial strains and growth conditions

Serotyping

Multi Locus Sequence Typing (MLST)

7F cps detection and sequencing

Cloning and *S. pneumoniae* transformation

Transmission Electron Microscopy (TEM)

Type 7F capsular polysaccharide detection by high-resolution magic angle

Spinning NMR spectroscopy (HR-MAS NMR)

Results

S. pneumoniae NT invasive strains AP422 and AP426 contain a 7F capsular locus with point mutations in *wchA*

Over-expression of a functional WchA restores the ability of *S. pneumoniae* to synthesize the 7F capsule

Discussion

References

Acknowledgements

Sommario

Streptococcus pneumoniae (*S. pneumoniae)* è uno tra i più importanti batteri patogeni per l'uomo se si considerano il tasso di mortalità e morbilità ad esso associati su scala mondiale (circa un milione e mezzo di morti ogni anno tra i bambini di età inferiore a 5 anni). *S. pneumoniae* è in grado di causare diverse malattie, sia non invasive che invasive, quali otite media, polmonite, sepsi e meningite, ma è anche annoverato tra i patogeni umani in grado di colonizzare asintomaticamente il tratto respiratorio di bambini e adulti sani. La colonizzazione è una fase necessaria per lo sviluppo di un'infezione respiratoria o sistemica ed è, infatti, un passaggio obbligato per il contagio nella comunità. I gruppi più a rischio d'infezione sono neonati, bambini al di sotto dei 5 anni di età, anziani ed individui immunocompromessi o immunodeficienti. Attualmente circa un milione di decessi l'anno vengono attribuiti ad infezioni da pneumoccocco e, di queste, il 50% circa coinvolgono bambini al di sotto dei 3 anni, o anziani affetti da polmonite. Il fallimento di trattamenti terapeutici antibiotici dovuto alla rapida diffusione di cloni contenenti elementi che conferiscono resistenza agli antibiotici anche di ultima generazione (quali i macrolidi), ha fatto sì che un enorme sforzo della comunità scientifica fosse diretto allo sviluppo di un vaccino in grado di prevenire l'insorgenza di patologie associate a *S. pneumoniae.*

S. pneumoniae è un batterio Gram-positivo caratterizzato dalla presenza sulla sua superficie di una spessa capsula polisaccaridica, considerata uno dei suoi principali fattori di virulenza. La capsula è, infatti, altamente immunogenica e in grado di indurre un aumento dell'invasività dello pneumococco favorendone la capacità di eludere i meccanismi di difesa aspecifici dell'ospite, primo tra tutti la fagocitosi. Ad oggi sono noti 94 sierotipi polisaccaridici capsulari differenti, identificati in base alle caratteristiche antigeniche del polisaccaride capsulare. Recenti studi di epidemiologia molecolare condotti in diverse regioni hanno dimostrato che i sierotipi capsulari hanno una distribuzione diversa a seconda della regione e che esiste una correlazione tra il sierotipo ed il grado di patogenicità. Inoltre, la presenza in individui sani di anticorpi specifici contro certi polisaccaridi capsulari garantisce la protezione in questi individui da infezioni causate da sierotipi omologhi.

Le precedenti osservazioni giustificano che tutti i vaccini attualmente in commercio siano a base polisaccaridica. In particolare, esistono vaccini polisaccaridici (Pneumovax 23) e polisaccaridici coniugati (Prevenar7, Prevenar13, Synflorix), tutti contenenti un numero limitato di sierotipi capsulari e in grado di indurre un'immunità sierotipo-specifica. Solo i vaccini a base di glicoconiugati sono immunogenici nei bambini al di sotto dei 2 anni di età, che, come detto precedentemente, costituiscono una delle maggiori categorie a rischio. Sebbene questi vaccini abbiano dimostrato una buona efficacia nel ridurre i casi di malattia invasiva causata da *S. pneumoniae*, il loro uso è limitato dal fatto che inducono un'immunità sierotipo-specifica e pertanto sono in grado di proteggere efficacemente solo contro le malattie invasive causate dai sierotipi contenuti nel vaccino stesso. Questo implica che, in base alla diversa

distribuzione dei sierotipi, i vaccini polisaccaridici in commercio possono avere un grado di copertura diverso a seconda della regione esaminata. Inoltre, la vaccinazione di massa in alcuni stati ha indotto una ridistribuzione dei sierotipi circolanti. Questo fenomeno, noto come "*serotype replacement*", consiste nello spiazzamento dei sierotipi contenuti nel vaccino da parte di altri non contenuti nel vaccino (*non-vaccine types*) e si pensa possa essere responsabile di una futura diminuzione della copertura garantita dai vaccini tuttora in commercio.

Per queste ragioni, il mondo della ricerca si è indirizzato verso lo sviluppo di un vaccino su base proteica che possa indurre una protezione sierotipo-indipendente contro l'infezione pneumococcica.

Un aspetto dibattuto della vaccinazione anti-pneumococco s'incentra sulla necessità che il vaccino debba prevenire la colonizzazione oppure semplicemente la malattia invasiva; è comunque chiaro che un vaccino ideale dovrebbe essere specie-specifico ed efficace nei gruppi a rischio.

L'identificazione di proteine immunogeniche conservate di pneumococco rappresenta quindi un passo cruciale. Infatti, le proteine identificate potrebbero essere utilizzate come proteine "*carrier*" in vaccini glicoconiugati, oppure come unici componenti di vaccini a base interamente proteica, se la protezione da esse fornita fosse adeguata. Questo ridurrebbe fortemente i costi di produzione, e renderebbe praticabile la distribuzione del vaccino anche nei paesi in via di sviluppo.

Tra le proteine selezionate nel corso degli anni come potenziali componenti di un vaccino vi sono la pneumolisina, PspA, PsaA, CbpA. Alla lista delle proteine storicamente già identificate come importanti antigeni recentemente si sono aggiunte le subunità proteiche che costituiscono il pilo-1, una struttura fibrillare identificata sulla superficie di alcuni ceppi di *S. pneumoniae.*

Come accade in molti altri batteri della specie streptococco, la superficie di *S. pneumoniae* è decorata da pili con caratteristiche adesive, composti da subunità legate covalentemente, importanti per la virulenza ed in particolare per l'adesione iniziale dei batteri alle cellule epiteliali.

In particolare, i pili dello pneumococco sono codificati da due regioni nel genoma denominate pilus islet-1 (PI-1) e pilus islet-2 (PI-2), presenti rispettivamente circa nel 30% e nel 16% dei ceppi di pneumococco circolanti. Sono note tre varianti di PI-1, clonalmente correlate, mentre esiste una sola variante di PI-2. La presenza delle due isole non correla con il sierotipo dei batteri, ma con il loro genotipo determinato tramite Multi Locus Sequence Typing (MLST). La prevalenza di ceppi positivi per la presenza di PI-1 o PI-2 è simile in isolati derivanti da malattie invasive o da *carriage*, mentre la prevalenza di ceppi positivi per PI-1 è più elevata in collezioni di ceppi antibiotico-resistenti, dove raggiunge anche il 60%. Quest'ultima osservazione favorisce l'ipotesi che, da un lato i batteri PI-1 positivi possano acquisire più

facilmente determinanti di resistenza, e, dall'altro, che l'espressione del pilo-1 possa costituire essa stessa un fattore di vantaggio per la diffusione della malattia.

Diversi studi hanno dimostrato che le subunità del pilo-1 (RrgA, RrgB ed RrgC) conferiscono un altissimo grado di protezione in modelli murini di infezione sia in esperimenti di immunizzazione attiva che passiva. Per questo motivo, nonostante il pilo-1 non sia presente in tutta la popolazione batterica e le sue subunità costituenti non siano altamente conservate, RrgA, RrgB ed RrgC sono considerati potenziali candidati per essere inseriti in un vaccino a base proteica di nuova generazione. RrgB in particolare, oltre ad essere la subunità più abbondante del pilo, essendo il costituente dello scheletro, è anche la più protettiva.

È nel contesto della caratterizzazione di fattori di virulenza di *S. pneumoniae* che si inserisce il mio lavoro di ricerca, incentratosi principalmente su due progetti, che mirano a: 1) definire se esiste un' associazione tra la presenza dell' isola codificante per il pilo-1 (PI-1) e frequenza di trasmissione di *S. pneumonie* e/o la durata della colonizzazione; 2) evidenziare il meccanismo molecolare associato all'inusuale assenza di espressione della capsula in ceppi isolati da malattia invasiva.

1. Assessment of *Streptococcus pneumoniae* PI-1 prevalence in carried and transmitted isolates from, mother-infant pairs on the Thailand-Burma border.

Dal momento che le proteine costituenti il pilo sono considerate potenziali candidate per un vaccino a base proteica contro lo *Streptococcus pneumoniae* e che il pilo-1 è considerato uno dei fattori importanti nell'iniziale adesione dello pneumococco alla cellula ospite durante la colonizzazione, lo scopo di questo lavoro è stato quello di valutare la prevalenza dell'isola del pilo-1 (PI-1) in una collezione di ceppi isolati da *carriage* durante uno studio longitudinale e di valutare la possibile associazione tra la presenza dell'isola stessa e la colonizzazione dell'ospite (valutata da parametri quali durata di colonizzazione e probabilità di trasmissione madre-figlio).

Lo studio ha previsto la caratterizzazione di ceppi colonizzatori del tratto nasofaringeo isolati tra il 2007 e il 2010 da 234 coppie di madri-bambini residenti nella zona di confine tra la Tailandia e la Birmania (zona in cui non è stato fino ad ora distribuito un vaccino contro lo pneumococco). Prelievi mensili alle madri e ai loro neonati, a partire dal primo mese di vita fino al compimento di un anno di età, hanno permesso di isolare 2496 ceppi di *Streptococcus pneumoniae*. Tutti i ceppi isolati sono stati sierotipizzati e si è ipotizzato che la presenza di uno stesso sierotipo in mesi successivi corrispondesse ad un singolo episodio di colonizzazione. Un sottogruppo di questi isolati (896) è stato poi scelto per essere rappresentativo degli episodi di colonizzazione e di trasmissione tra madre e figlio verificatisi durante il periodo di osservazione e i cui criteri di definizione sono stati stabiliti all'interno dello

studio. Questi isolati, rappresentativi di 34 sierotipi, sono stati caratterizzati dal punto di vista genotipico mediante MLST (Multi Locus Sequence Typing) e la presenza di PI-1 valutata tramite PCR. La procedura MLST prevede l'amplificazione e il sequenziamento di 7 frammenti appartenenti a 7 geni costitutivamente espressi di pneumococco. La determinazione della loro sequenza consente l'assegnazione di un numero di allele per ogni gene, mentre la combinazione dei 7 alleli permette la determinazione univoca di un numero, denominato *Sequence Type* (ST).

Degli 896 ceppi analizzati per 887 è stata determinata con successo la positività per PI-1 che è risultata essere pari al 35.2% (312/887). In particolare, PI-1 è associata soltanto a 10 dei 34 sierotipi (in misura maggiore ai sierotipi 19F e 23F) a cui appartengono i ceppi analizzati.

Un'analisi più dettagliata ha permesso poi di valutare la possibile associazione tra la presenza e/o assenza di PI-1 e la trasmissione del patogeno tra madri e rispettivi figli. Un pannello ristretto di ceppi (483 isolati), comprendenti sia isolati trasmessi che non trasmessi è stato selezionato per tale scopo (l'effettiva presenza di casi di trasmissione è stata determinata dalla concordanza di sierotipo e ST in ceppi isolati da coppie madre-figlio). In particolare, nell'analisi sono stai inclusi i ceppi appartenenti ai sierotipi presenti nel vaccino tredici-valente (Prevnar-13) e i ceppi non sierotipizzabili, che costituiscono una frazione molto abbondante dell'intera collezione.

Dei 483 ceppi analizzati, 219 sono risultati essere positivi per la presenza di PI-1 (45.3%) con una frequenza molto più elevata per sierotipi come 19F e 23F dove la presenza dell'isola è stata trovata rispettivamente nel 91% e nel 77% dei ceppi. L'associazione tra sierotipo e presenza del pilo osservata in questo studio va ricercata soprattutto nel genotipo di appartenenza dei ceppi analizzati e può essere spiegata con la differente diffusione regionale che molto spesso si osserva per infezioni/colonizzazioni da parte dello pneumococco; inoltre, tale associazione può dipendere dai criteri di selezione applicati allo studio.

L'analisi della trasmissibilità dei ceppi tra madri e rispettivi neonati non ha evidenziato alcuna correlazione tra la presenza dell'isola del pilo e la trasmissione del ceppo di *S. pneumoniae*. Infatti, non c'è differenza statisticamente significativa tra la prevalenza di PI-1 in ceppi trasmessi (47.5%) tra madre e figlio rispetto a quelli non trasmessi (43.3%) (OR1.2; 95% CI: 0.8: p=0.4).

Inoltre, per quanto riguarda l'analisi dell'associazione tra la presenza di PI-1 e la durata del periodo di colonizzazione, gli episodi di *carriage* sono risultati essere complessivamente più lunghi per gli isolati PI-1 positivi. Tuttavia, questo dato perde di significatività statistica effettuando l'analisi tenendo conto di un elemento confondente quale il sierotipo (essendo noto che tipi capsulari diversi sono associati a durate di *carriage* diverse, visto il piccolo numero di isolati analizzati e l'associazione PI-1/clone/sierotipo).

In conclusione, la collezione analizzata, pur provenendo da un'area geografica (Sud-est Asia) epidemiologicamente ancora non ben caratterizzata, presenta una prevalenza di ceppi PI-1 positivi in linea con precedenti studi epidemiologici. Inoltre, l'analisi della collezione non ha evidenziato particolari associazioni tra la presenza di PI-1 e la trasmissibilità e/o la durata temporale del *carriage*.

1. Point mutations in wchA are responsible for non-typeability of two invasive Streptococcus pneumoniae isolates.

La capsula polisaccaridica è uno dei fattori di virulenza più importanti per lo pneumococco e la sua espressione è strettamente associata all'invasività. La capsula conferisce infatti agli streptococchi una maggiore resistenza alla fagocitosi mediata dalle cellule dell'ospite e garantisce quindi la sua persistenza in distretti quali l'apparato circolatorio.

Sebbene, come precedentemente indicato, gli isolati di *S. pneumoniae* presentino generalmente una capsula sulla loro superficie, esiste una piccola percentuale di ceppi di *S. pneumoniae*, che non presenta una reazione positiva con nessuno dei sieri usati per la tipizzazione. I ceppi di questa sotto-categoria sono stati definiti Non-Typeable (cioè non tipizzabili, NT). I ceppi NT o sono acapsulati o esprimono una quantità di capsula molto limitata, non sufficiente per essere evidenziata con le tradizionali tecniche di rilevamento. Per questo motivo, i ceppi NT sono isolati quasi esclusivamente da episodi (epidemie) di congiuntivite o da tamponi naso-faringei e sono ritrovati solo molto raramente in isolamenti effettuati da infezioni invasive dove l'assenza della capsula rende la loro persistenza piuttosto improbabile.

Le basi genetiche della non tipizzazione, che alterano l'espressione di una o più proteine del locus capsulare, sono state studiate in generale solo per collezioni piccole e scarsamente rappresentative e in generale per ceppi NT isolati da casi di colonizzazione o congiuntivite. I meccanismi molecolari evidenziati finora come responsabili della non tipizzazione sono la perdita di parte o dell'intero locus capsulare (contenente i geni che codificano per le proteine responsabili dell'assemblaggio dei polisaccaridi), l'inserzione di geni, mutazioni puntiformi o duplicazioni di sequenza.

L'analisi genetica di ceppi NT (tramite tipizzazione MLST, vedi capitolo 1) ha evidenziato che essi possono appartenere a cloni tipicamente associati a NT pneumococci o, più raramente, a cloni generalmente associati a sierotipi.

Visto le sopra elencate premesse, l'isolamento nel 2006 di ceppi NT da due casi di polmonite bacteriemica nello stesso ospedale italiano è stato considerato un evento raro. Per questo motivo si è deciso di procedere alla caratterizzazione genetica della loro non-tipizzabilità.

La genotipizzazione dei due isolati tramite MLST ha evidenziato che entrambi appartengono all'ST191 che è generalmente associato al sierotipo 7F. Successivamente, il sequenziamento dell'intero locus capsulare dei due ceppi non-tipizzabili ha confermato la presenza dell'intero locus capsulare di tipo 7F. Un'analisi dettagliata della sequenza ha consentito di evidenziare in ognuno dei due ceppi la presenza di una mutazione puntiforme nel gene *wchA* (mutazioni diverse nei due ceppi), codificante per la proteina glicosiltransferasi che, in generale, è responsabile del primo passaggio di assemblaggio del polisaccaride. Entrambe le mutazioni comportano la traduzione di una proteina WchA tronca all'estremità C-terminale, che risulta probabilmente non funzionale.

Al fine, poi, di verificare l'effettiva responsabilità delle due mutazioni presenti nel gene *wchA* della non-tipizzabilità dei ceppi testati, si sono trasformati entrambi gli isolati non-tipizzabili con un plasmide pMU1328 contenente un gene *7F wchA* codificante per una proteina WchA funzionale.

I due ceppi over-esprimenti WchA e i due ceppi non-tipizzabili sono stati analizzati mediante spettroscopia NMR e microscopia elettronica al fine di verificarne il fenotipo e misurare lo spessore di un'eventuale capsula presente sulla superficie batterica. E' stata confermata l'assenza di una capsula polisaccaridica sulla superficie dei ceppi non-tipizzabili, mentre il ripristino di WchA è stato sufficiente a riattivare il normale assemblaggio di una capsula 7F in entrambi i ceppi NT.

In conclusione, in questo studio si sono identificate due nuove mutazioni puntiformi spontanee responsabili della non-tipizzabilità di pneumococco e si è inoltre dimostrato che WchA è essenziale per una corretta biosintesi del tipo capsulare 7F. Tuttavia, assunto che i due NTPn analizzati non sono in grado di assemblare una capsula a causa di una mutazione puntiforme incorsa in un gene essenziale, resta da chiarire come i due ceppi siano riusciti a causare una malattia invasiva. La mancanza di dati clinici riguardo ai pazienti non ha permesso la formulazione di una teoria a tale proposito.

Introduction

Introduction

1. A brief history of vaccines.

Vaccines in the era of genomics.

Vaccination, together with the wider availability of potable water, has had the most profound positive effect on the quality of the public health. During the past century, vaccination essentially eliminated many infectious diseases causing mortality in infants and in children. Less than two hundred years ago life expectancy was around 40 years of age, a condition that still persists in some undeveloped countries. From about the middle of the nineteenth century, great improvements in living standards were made in industrialized countries, initially by providing adequate sanitation facilities and safe drinking water. However, this was not enough to prevent epidemics and sometimes pandemics of infectious diseases. In fact, the “Spanish flu” in 1918-19 killed over 30 million people, more than all those who died in World War I during 1914-1918; and Smallpox was a terrible killer for centuries, responsible for many millions of victims. Coping with these issues became progressively possible only through the development of vaccines against some of the major diseases.

Vaccines against diphtheria, tetanus, polio, measles, mumps, rubella, pneumococcus, hepatitis B and meningitis have reduced the incidence and mortality of these diseases by > 97-99% (Rappuoli *et al.*, 2002). Nevertheless, perception of vaccines in the public opinion is not completely positive. A lot of people are still skeptical about the real need of vaccines, and this attitude has been particularly evident during the influenza A (H1N1) pandemic in 2009. Indeed, persons tend not to realize the importance of illness prevention, if the disease is not causing the death of the patients. Novel vaccines in fact do not only prevent deaths directly caused by pathogens but can also prevent malignancies associated with them (e.g. papillomavirus, hepatitis B virus and *Helicobacter pylori*). Prevention of infections can increase patient fitness (by decreasing the number of infectious episodes and associated sequelae), and, reducing the number of harmful inflammatory processes happening during infections, prolong life expectancy (Crimmins and Finch, 2006). In addition, the socio-economic costs of hospitalizations and the positive effects associated with extensive vaccination programs, like the overall reduction of the incidence of infections (herd immunity) and the reduced use of antibiotics and antivirals (decreasing the emergence of resistances), are not always duly taken into account.

Infectious diseases, despite antimicrobial therapies, remain the leading cause of death worldwide. In addition, infectious diseases for which a cure is not available or not efficacious are expected to emerge (or re-emerge) in the coming decades. Therefore, the development of novel vaccines against these diseases is mandatory (Jones *et al.*, 2008; Rappuoli, 2004).

The majority of the vaccines available to date (e.g. polio, smallpox, pertussis and tetanus) were developed using conventional vaccinology strategies based on the Pasteur’s principle of “isolate, inactivate and inject” (Rappuoli 2004; Robbins *et al.*, 2005). Although very successful, this approach failed against certain pathogens and, in other occasions, the obtained vaccines are no longer adequate due to safety concerns and low efficacy.

In detail, the conventional approach has been always successful in developing vaccines against pathogens that do not change the vaccine-targeted antigens over time (and with an overall low degree of species variability) and for which protection is

antibody-mediated (Rappuoli, 2007). However, it has been also successful in targeting particular pathogens displaying many variants of target-vaccine antigens. Examples of such vaccines are the ones based on polysaccharide antigens of pneumococcus and meningococcus. On the other hand, as already mentioned above, for certain pathogens the conventional vaccine development approach has failed. The reasons for these failures have been probably the high antigen variability and the fact that protection is mediated by mechanisms which are still not clearly understood. For all these cases, a huge investment to implement new methods for novel vaccine candidate discovery and to develop new technologies to express, deliver and formulate antigens, is necessary.

A modern vaccine design: The Reverse Vaccinology.

In the last two decades, important discoveries have enlightened the science of molecular immunology, proteomics, physicochemestry, genomics, and systems biology. These advances have totally changed the way vaccines are designed, characterized and developed.
Genome sequencing has become routine, and modern vaccine design is taking advantage of the accumulating genomic information, revolutionizing the study of microbial pathogenesis, genetics and population biology. The alliance between technological innovations, such as high-throughput sequencing, coupled with reverse genomics and conventional vaccinology has led to a new paradigm of vaccine development (Scarselli *et al.*, 2005).
In the year 2000, a revolutionary approach to identify and develop new vaccines was published (Pizza *et al.*, 2000). The strategy was called reverse vaccinology (RV), since the initial identification of vaccine candidates was not performed based on experimental observations but on a bioinformatic analysis of the pathogens' genome. In detail, RV is based on a "multiple antigen selection approach" comprising in *silico* antigen prediction, functional genomics and comparative genomics; the final identification of the vaccine candidates is then based on epidemiological data (conservation and expression of the antigens) and immunogenicity/protection screenings, performed with animal model experiments, human sera analysis, and expression libraries screenings (Fig. 1).

This approach has largely replaced classical vaccinology methods. Its main advantage is the fast *in silico* prediction of vaccine candidates, which will mainly be surface exposed proteins, most likely accessible to antibodies binding during infection (Rappuoli, 2004). Interestingly, another advantage of this approach is that it could be applied to non-cultivable microorganism, difficult or impossible to target with conventional approaches.

The RV methodology has been initially applied on the serogroup B *Neisseria meningitidis* (MenB). The project served as proof of concept of the validity of the method and led to the development of a multivalent recombinant vaccine that is now in phase III clinical trials. The MenB project demonstrated the strength, the speed and all the advantages of the RV over classical vaccine approaches.

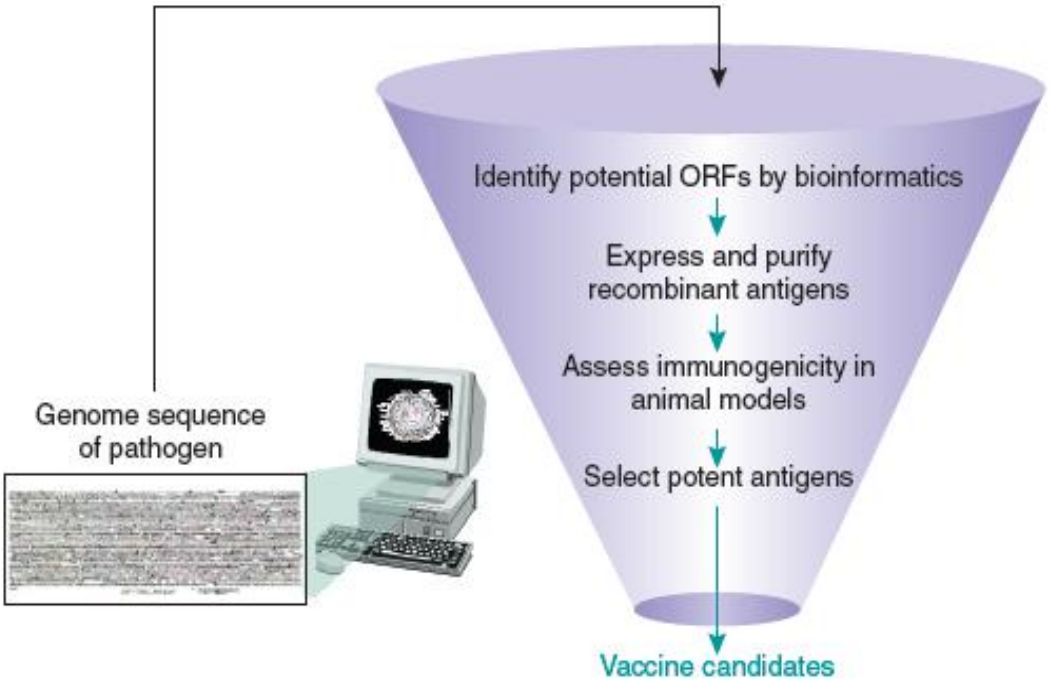


Fig. 1: The complete reverse vaccinology process. Reverse vaccinology starts with antigen prediction based on *in silico* analysis as well as comparative genomics and functional genomics. The availability of genomes and bioinformatic tools is essential to the *in silico* prediction of all potential open reading frames (ORFs) as well as comparative and functional genomics and proteomics. During the next phase, the antigens are purified and screened for protection efficacy, epidemiological relevance, expression profile, immunoreactivity against human sera. They are also characterized for their function and toxicity. Ideally, a vaccine should consist of one to five different antigens.

2. Epidemiological studies: aims and relevance.

Definition and objectives of Epidemiology.

Epidemiology is the study of how disease is distributed in populations and of the factors that influence or determine this distribution. Why does a disease develop in some people and not in others? The premise underlying epidemiology is that disease, illness, and ill health are not randomly distributed in a population. Rather, each of us has certain characteristics that predispose to, or protect against, a variety of different diseases. These characteristics may be primarily genetic or the result of exposures to certain environmental hazards (Fig. 2).

More broadly epidemiology is defined as:

"The study of the distribution and determinants of health-related states or events in a specified population and the application of this study to control of health problems" (Last JM: A dictionary of Epidemiology, ed. 2 New York, Oxford University Press, 1988)

Noteworthy, this definition incorporates both a description of the content of the discipline and the purpose or application for which epidemiologic investigations are carried out.

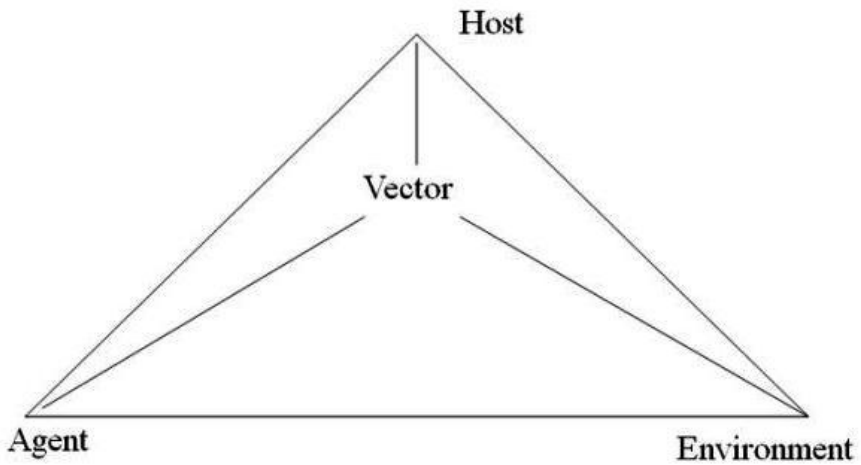


Fig. 2: The epidemiologic triad of a disease. According to the diagram, a disease is the product of an interaction of the human host, an infectious or other type of agent and the environment that promotes the exposure.

What are the specific objectives of epidemiology?

First, to identify the etiology or the cause of a disease and the risk factors-that is, factors that increase a person's risk for a disease.

It is important to understand how the disease is transmitted from one person to another or from a non-human reservoir to a human population. The ultimate aim is, in fact, to intervene to reduce morbidity and mortality caused by the disease. The identification of the etiologic or causal factors of the disease and the reduction or elimination of the exposure to those factors, are the basis for the development of rationally designed prevention programs.

Second, to determine the extent of disease found in the community.

What is the burden of disease in the community? This question is critical for planning health services and facilities and for training future health care providers.

Third, to study the natural history and prognosis of disease.

Clearly, certain diseases are more severe than others; some may be rapidly lethal, and others may have longer or shorter durations of survival. The baseline natural history of a disease has to be defined in quantitative terms to develop new modes of intervention, either through treatments or new ways of preventing complications.

Fourth, to evaluate new preventive and therapeutic measures and new modes of health care delivery.

The development of effective vaccines specifically targeting the major invasive pathogens represents a significant advance in improving the public health.

Fifth, to provide the foundation for developing public policy and regulatory decisions related to environmental problems.

The Molecular Epidemiology of infectious disease: a simple extension of the Epidemiology definition.

Epidemiology, as said before, is often defined as:

"*the study of the distribution and determinants of diseases and injuries in human populations*" (Mausner *et al.*, 1974).

A simple extension of this definition to molecular epidemiology of infectious diseases is:

"the study of the distribution and determinants of infectious diseases that utilizes molecular biology methods."

As early as 1977, Higginson defined molecular epidemiology as:

"*the application of sophisticated techniques to the epidemiologic study of biological material*" (Higginson, 1977).

More recently, Levin et al. stated:

"*The practical goals of molecular epidemiology are to identify the microparasites responsible for infectious diseases and determine their physical sources, their biological relationships, and their route of transmission and those of the genes responsible for their virulence, vaccine-relevant antigens and drug resistance*" (Levin *et al.*, 1999).

In general, the word "determinants" in the definition of epidemiology is emphasized, as the main goal of nearly all molecular epidemiologic studies is the characterization of determinants of disease distribution.

It is recognized that infectious disease distribution and determinants of this distribution may be identified by molecular biology techniques; it is also becoming evident that infectious disease distribution in a community of hosts may itself be determined or "programmed" by a pathogen's own genetic material that evolved in response to hosts' antimicrobial machinery as well as to the environment in which the pathogen resides. That is, every human, animal, and plant pathogen has evolved to transmit itself to a host or to a community of hosts in its own characteristic fashion within a specific environmental context. Mechanisms that underlie an organism's specific pattern of transmission may, thus, be genetically and environmentally determined. The discipline of molecular epidemiology, therefore, should also include the study of genetic factors that determine and regulate an organism's specific pattern of transmission.

The relevance of Epidemiological Studies.

In recent years, epidemiology has become an essential approach in both health and clinical practice, while prevention and therapy are still too often viewed as mutually exclusive activities. In fact, this discipline is the basic science for the development of disease prevention strategies, playing major roles in the development and evaluation of public health policies. Epidemiology is used together with laboratory research to identify risk factors for disease and to shed light on the mechanisms involved in pathogenesis. Indeed, epidemiologic studies play a fundamental role in the entire process of drug or vaccine development during all of the phases: pre-clinical, clinical and post-market.

Actually, one of the main objectives of the public health community consists in the reduction of infectious disease burden due to different invasive pathogens, among which the pneumococcus is one of the most important. In fact, *Streptococcus pneumoniae* causes high morbidity and mortality throughout the world (even in regions where antibiotics are readily available) and, together with HIV, malaria, and tubercolosis, it represents one of the four major infectious disease killers.

3. Streptococcus pneumoniae.

Pathogenesis of pneumococcal disease: colonization and invasion.

The Gram-positive bacterium *Streptococcus pneumonia* (*S. pneumoniae*), also called pneumococcus, was first isolated by Pasteur in 1881 from the saliva of a patient with rabies. Typically, it is observed in pairs (diplococci) but sometimes also single bacteria or in short chains (Fig. 3).

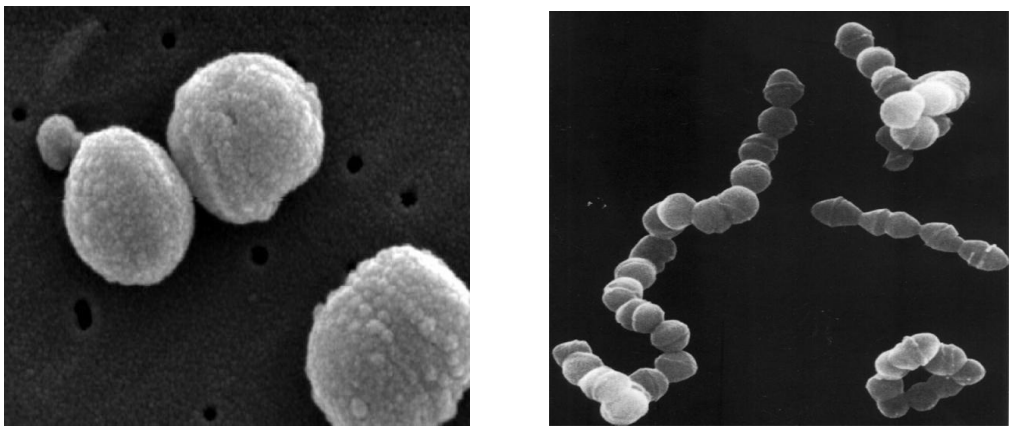


Fig. 3: General representation of the bacterium *Streptococcus pneumoniae*.

S. pneumoniae bacteria are part of the human nasopharyngeal commensal flora but are also able to cause a variety of illnesses, which include invasive pneumococcal diseases (IPDs) such as bacteremia, meningitis and bacteremic pneumonia. Furthermore, pneumococci are a common cause of mucosal infections such as otitis media, sinusitis, mastoiditis and non-bacteremic pneumonia (Rose *et al.*, 2009) (Fig. 4).

Colonization is commonly followed by horizontal dissemination of the pathogens to individuals in the direct environment, thus leading to spread within the community (Givon-Lavi *et al.*, 2002). The reported rates of bacterial acquisition and carriage depend on age, geographical area, genetic background, and socioeconomic conditions; risk groups for diseases caused by pneumococci include children under the age of 2 years, elderly and patients with immunodeficiencies. *S. pneumoniae* colonization is a dynamic process; indeed, the competition with the resident flora for the persistence in the human airways has an important role in determining the colonization rates of pneumococcus and other virulent microorganisms. Besides, also the efficiency of the local host immune response, strictly depending on the condition of the upper respiratory tract, has to be taken into account; in fact, a poor mucosal immune response can lead to persistent and recurrent colonization and consequent dissemination of the infection, whereas a brisk local immune response to the pathogen can eliminate colonization and prevent re-colonization.

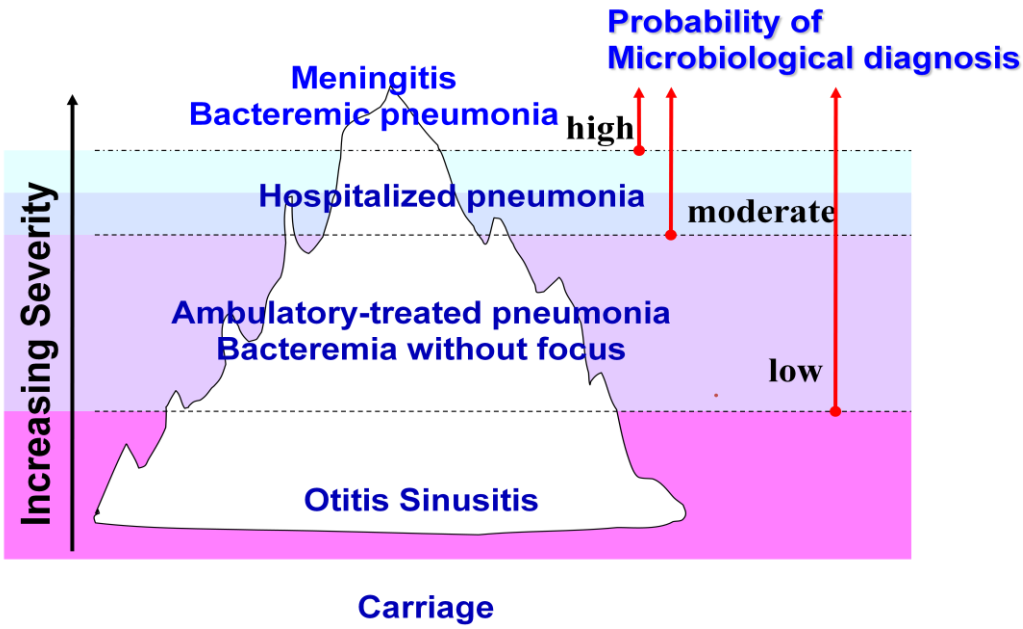


Fig. 4: Diagram of the different pneumococcal diseases correlating with the probability of microbiology diagnosis. For example, it's very hard to recover samples of the s pneumonia strains that cause bacteremic pneumoniae.

This re-colonization process is very common, and, in fact, probably all humans are colonized with this organism at least once early in their life. Especially in crowded environments, such as day-care centers, nursing homes, hospitals and jails, the risk of colonization is extremely high. As already mentioned before, colonization is not usually followed by disease, since this is prevented by the innate and adaptive immune system. However, disturbance of homeostasis between host and pathogen, for example through viral infections, malnutrition or local damage of the mucosa, is associated with an increased development of pneumococcal (invasive) diseases.
It is generally accepted that *Streptococcus pneumoniae* infection can progress to respiratory or even systemic disease following the initial, mostly symptomless, colonization of the nasopharynx (Bogaert *et al.*, 2004). However, despite the high burden of disease, the pathogenic mechanisms exploited by *S. pneumoniae* to

adhere to human epithelia, most likely triggering the colonization process, and to then progress to invasive disease are still poorly understood.
A major virulence factor of *Streptococcus pneumoniae* is the polysaccharide capsule. The pneumococcal capsule displays an extremely high variability (it exists in fact in at least ninety different types, known as serotypes) and provides an effective barrier against host-cell mediated phagocytosis allowing bacterial persistence in the blood. In addition to the capsular determinant, many other secreted or surface exposed factors have been described to be of importance for virulence through *in vivo* animal model studies and *in vitro* experiments (i.e. choline-binding protein A (CbpA), the pneumococcal toxin pneumolysin, pilus components, pneumococcal surface protein A (PspA), pneumococcal surface adhesin A (PsaA), pilus components); however, their direct contribution to and essentiality for disease development in humans have still to be determined.

Pneumococcal vaccines: polysaccharides and surface proteins

Streptococcus pneumoniae-related infections have a major global impact on healthcare, especially in the developing world, and are considered the number one vaccine-preventable cause of death in children (Adam D., 2009). In fact, *S. pneumoniae* is the single commonest cause of community-acquired pneumonia, and has become the most frequent cause of meningitis in many regions. Although determining the exact burden of pneumococcal disease is complicated by difficulties in establishing an aetiological diagnosis, particularly in cases of non-bacteremic pneumonia, the pneumococcus is conservatively estimated to kill more than a million children under the age of 5 years each year in developing countries (Fig. 5). This accounts for 20-25% of all deaths in this age group (Williams *et al.*, 2002; WHO, 2007). Even in developed countries, where effective antimicrobial drugs are readily available, morbidity and mortality from pneumococcal disease is significant. The other major driver for vaccine development has been the increasing threat posed by antibiotic-resistant pneumococci. Since the discovery of the antibacterial properties of penicillin by Fleming in 1929, many antibiotics have been successfully used for treatment of pneumococcal infections. However, recently, antibiotic resistance has become a worldwide problem, limiting the choice of antimicrobial agents. Therefore, prevention of pneumococcal disease through vaccination has become of even greater interest.
The increasing prevalence of penicillin- and multi-drug resistant pneumococci makes difficult the management of patients with suspected pneumococcal disease, particularly those with meningitis. In developed countries the problem is at a certain extent overcome by the use of more expensive alternative antimicrobials, but this option is not available in poorer parts of the world.
All the vaccines that are currently present on the market are based on a limited number of pneumococcal capsular polysaccharides, conjugated or not to a carrier protein (from 7 to 13 different capsular types for the conjugated and 23 capsular polysaccharides for the unconjugated vaccines). These vaccines have all demonstrated good efficacy in reducing invasive disease caused by the serotypes included in the vaccines, and a limited efficacy against acute otitis media. In addition, the 7-valent, 10-valent and 13-valent conjugate vaccines, unlike the 23-valent polysaccharide non-conjugated vaccine, have proved good effectiveness in children under 2 years of age. However, due to the extremely high variability of the

pneumococcal capsule composition, and the different geographical distribution of the vaccine pneumococcal serotypes, the efficacy of these vaccines is mostly limited by their serotype-dependent coverage. In addition, following the vaccine introduction in the medical practice, epidemiological surveillance studies have reported an increase in the disease caused by non-vaccine types (known as serotype replacement events).

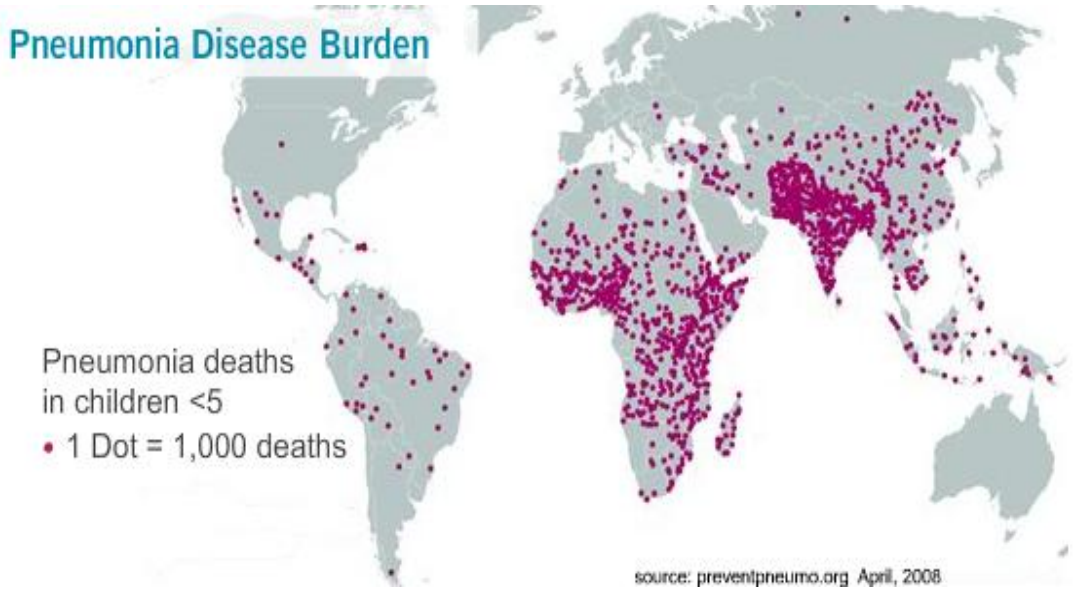


Fig. 5: Graphic representation of pneumonia disease burden in children under the age of 5 years each year, based on WHO estimates.

Moreover, the prohibitive cost of producing conjugate vaccines is a limit for their use in developing countries with the highest burden of disease and poor economical resources. For these reasons, vaccine research is currently focusing toward the use of one or more universally conserved protein antigens as potential vaccine components, able to elicit serotype-independent protection (Ryan and Antonelli, 2000). In this respect, cell-surface proteins, implicated in the interaction with host cells and tissues and playing an important role in pneumococcal virulence (Navarre and Schneewind, 1999), have attracted considerable attention and are currently under investigation as possible components of a multi-valent protein–based vaccine targeting pneumococcal disease (i.e. PcsB, PhtD, Stkp, PspA, RrgB pilus backbone, PspC, pneumolysin).
Indeed, pneumococcal proteins can be exposed on the bacterial surface through different mechanisms; on this basis they can be distinguished in membrane (possessing one or more trans-membrane domains), cell-wall anchored (linked to the cell-wall through an LPXTG motif), lipoproteins (covalently linked to a phospholipid) or choline binding (bound to choline residues decorating teichoic acid (TA) and lipoteichoic) acid proteins.

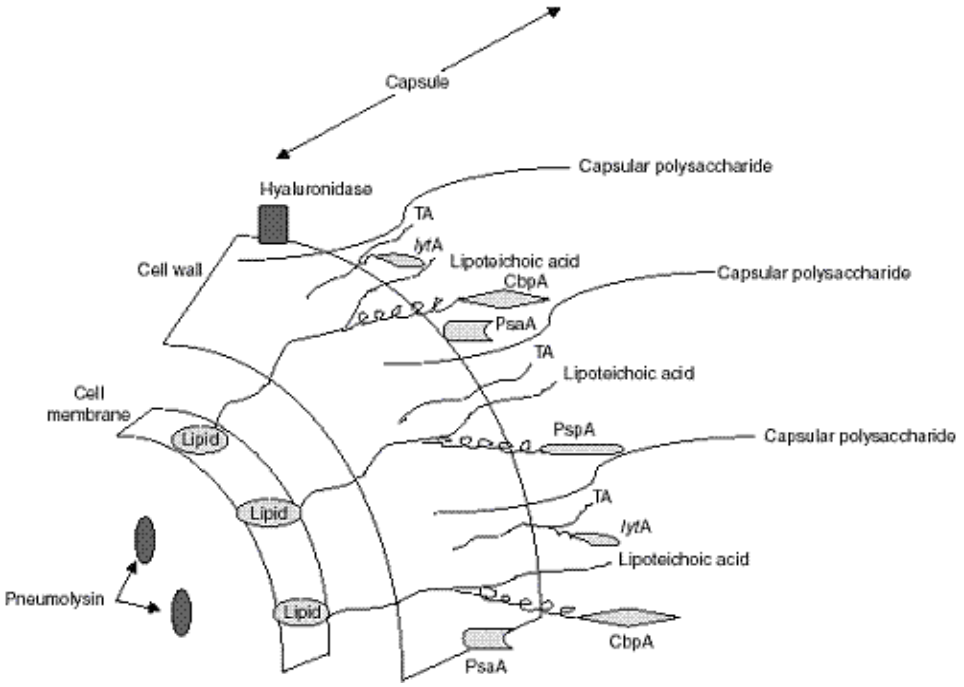


Fig. 6: Diagram of pneumococcal surface. Cell wall anchored proteins (CWAPs) are represented as well as some of the known virulence factor.

Briefly, cell wall anchored proteins (CWAPs) are attached to the bacterial surface through a covalent linkage with the peptidoglycan (Fig. 6). Formation of this linkage is catalyzed by specialized membrane-associated transpeptidases known as sortases, which recognize the C-terminal LPxTG motif of proteins destined to be attached to the cell wall (Ton-That *et al.*,2004a). Immunoglobulin A1 protease, hyaluronidase and pilus components are examples of pneumococcal proteins known to be involved in bacterial virulence belonging to this category. Besides*, S. pneumoniae* is the only human-pathogenic bacterium thus far known to express surface proteins specifically binding to choline as a mechanism of surface attachment. Interestingly, several of the choline binding proteins (Cbps), like CbpA, CbpD, CbpE and CbpG as well as all three pneumococcal autolysins (LytA, LytB and LytC) and also PspA (all belonging to this category) are described as bacterial factors implicated in the colonization processes.

Epidemiology of the pneumococcus: serotyping vs. molecular typing.

More than 90 serotypes have been identified based on the antigenic composition of the polysaccharide capsule. All the serotypes exhibit a distinctive epidemiology with regard to their potential to cause invasive disease, their occurrence in specific age groups or geographic regions, their association with antibiotic resistance and their epidemic potential (Hausedorff et al., 2005), thus accounting for the serotype dependent vaccine coverage described above. Indeed, most *S. pneumoniae* serotypes have been associated at various degrees to invasive diseases, but, in each region, only few serotypes (and mainly region specific) account for the majority of pneumococcal infections, with serotype composition changing upon seasonal and epidemic variations due to different clones spread. Serotypes 1, 5, 6, 7, 14, 18, 19,

and 23 are the most prevalent worldwide, accounting for 60-80% of the invasive diseases, depending on the area of the world. On the other hand, the serotype distribution among nasopharyngeal carriage isolates varies considerably by country, age-group, and type of cohort analyzed.
Capsular serotyping is currently the most widely used and efficacious strategy for classification of pneumococcal strains; however, this method presents some limitations like the inability to quantify the genetic relationship between isolates (population genetic) and the different accuracy in determining the serotypes achieved by different laboratories. In order to overcome these classification limits, a nucleotide sequence based method, Multi Locus Sequence Typing (MLST), has been developed with a global epidemiology perspective. This method involves the sequencing of internal fragments from seven house-keeping genes, and leads to the unequivocal assignment of the isolates to a sequence type (ST) (Chan et al., 2001). The strength of this approach is that sequence data are unambiguous, can be held on a central database (http://www.mlst.net) and be queried through a web server. On the other side, the expensiveness of this gene based strategy has strongly limited the diffusion of MLST typing to western rich countries. Interestingly, there is no univocal correspondence between the two classification methods, therefore, in defining an isolates both serotyping and MLST typing are needed, as serotypes are usually associated to more than an ST type and vice-versa. (Chan et al., 2001).
Notably, the increased use of combined MLST and serotyping classifications has allowed following the spread through different regions and during time of specific clones, often “disguised” with different capsular serotypes. The genetic fitness factors, other than antibiotic resistance, responsible for the successful spread of these clones are currently under investigation.

The pneumococcal capsule and its biosynthesis.

As already mentioned above, the polysaccharide capsule is the most important virulence factor of the pneumococcus as it protects the bacteria from phagocytosis. Factor (typing) sera are used to divide pneumococci into serotypes and serogroups, which include immunologically related serotypes. At present, 92 individual serotypes are recognized by their patterns of reactivity with the factor sera (Bentley *et al.*, 2006; Mavroidi *et al.*, 2007a; Park *et al.*, 2007), and serotypes vary in the extent to which they are carried in the nasopharynx and the degree to which they are recovered from different disease states (Hausdroff *et al.*, 2000; Spratt *et al.*, 2004). Expression of a capsule is important for survival in the blood and is strongly associated with the ability of pneumococci to cause invasive disease. The capsule is surface exposed, and antibodies against CPS provide a type-specific protection against pneumococcal disease. In detail, the capsule consists of high-molecular-weight polymers made up of units of repeated oligosaccharides (OS) which can contain between two and eight monosaccharides. Many serotypes possess acidic components (like D-glucuronic acid or phosphate groups), ribitol, or arabinitol (Van Dam *et al.*, 1990). In six serotypes, phoshporylcholine (PC) is part of capsular PS (Sørensen *et al.*, 1984).
With the exception of types 3 and 37, which are synthesized by the synthase pathway (Llull *et al., 1999;* Cartee *et al.,* 2001; Dillard *et al.,* 1995; Arrecubieta *et al.,* 1994; Waite *et al.,* 2003), pneumococcal CPSs are generally synthesized by the Wzx/Wzy-dependent pathway (Fig. 7). The genes for the latter pathway are located at the same chromosomal locus (*cps*), between *dexB* and *aliA* (Yother J., 2004;

Kolkman *et al.,* 1998; Garcia *et al.,* 2000). CPSs are synthesized by transfer of an initial monosaccharide phosphate from a nucleotide diphosphate sugar to a membrane-associated lipid carrier, followed by the sequential transfer of further monosaccharides to produce the lipid-linked repeat unit. This is transferred to the outer face of the cytoplasmic membrane by the repeat-unit transporter or flippase, polymerized to form the mature CPS, and then attached to the peptidoglycan (Sorensen *et al.*, 1990). The *cps* locus therefore typically encodes the enzymes necessary to build the repeat unit, (including an initial glycosyl phosphate transferase, and additional transferases responsible for the formation of the linkages), to allow the addition of sugars (or other moieties), or otherwise modify the repeat unit, as well as the repeat-unit flippase and polymerase (Yother J., 2004). The substantial diversity of pneumococcal CPSs is believed to have arisen as a consequence of selection for antigenic diversity imposed by the human immune system (Spratt *et al.*, 2004). The evolutionary timescales and the genetic events by which novel serogroups and serotypes arise are unclear. Comparisons of the available *cps* loci indicate a variety of genetic mechanisms and show that the central genes responsible for the synthesis and polymerization of the repeat unit are highly variable and often non-homologous between serotypes. These genes have a low percentage G+C content, and new serotypes may frequently have been generated by the introduction of novel *cps* genes into pneumococci by lateral gene transfer from other species. A much better understanding of the complex mechanisms by which antigenic diversity arises could be obtained by analyzing the sequences of the complete set of pneumococcal *cps* loci (Bentley *et al.*, 2006).

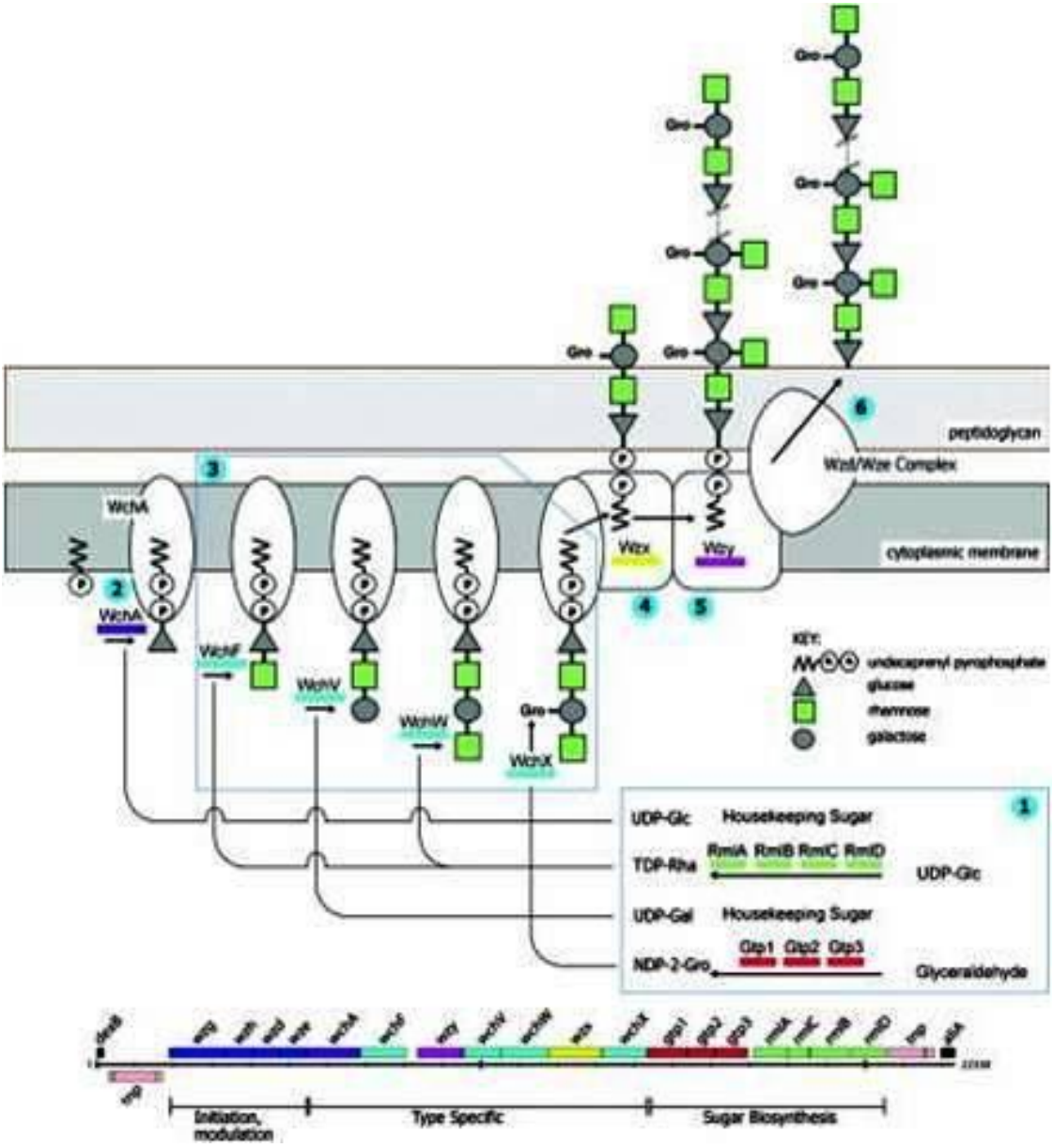


Fig. 7: Representation of the Wzx/Wzy-Dependent Pathway for Biosynthesis of CPS 9A. Pictured is a hypothetical model for capsule biosynthesis in *S. pneumoniae* based on a mixture of experimental evidence and speculation. For a recent review, see Yother 2004.
1) Non-housekeeping nucleotide sugar biosynthesis.
2) The initial transferase (WchA in this case) links the initial sugar as a sugar phosphate (Glc-P) to a membrane-associated lipid carrier (widely assumed to be undecaprenyl phosphate).
3) Glycosyl transferases sequentially link further sugars to generate repeat unit.
4) Wzx flippase transports the repeat unit across the cytoplesmic membrane.
5) Wzy polymerase links individual repeat units to form lipid-linked CPS.
6) Wzd/Wze complex translocates mature CPS to the cell surface and may be responsible for the attachment to the peptidoglycan. The complex of WchA, Wzy, Wzx, Wzd and Wze shown in the membrane is based on that in Figure 2 of Whitfield and Paiment (ref inserire) for the related *Escherichia coli* Type 1 capsule.

Pili in *Streptococcus pneumonia*e.

The surface-exposed components of a bacterium permit the microorganism to assess its environment; hence, these constituents often are major virulence factors. Indeed, pili and fimbriae are long filamentous structures that are presented by bacteria on their surface. They play a key role in host cell invasion, biofilm formation, cell aggregation, DNA transfer and twitching motility as it has been reported in different studies (Proft and Baker, 2009).
The role of pili as adhesive organelles is of particular importance for pathogenic bacteria; in fact, they are involved in roles such as attachment to specific host cells during colonization, a key step in establishing an infection.
It has been suggested that bacteria use these structures to form an initial association with host cells, which can then be followed by a more “intimate” attachment that brings the bacterium into close proximity to the host-cell surface.
Pili of Gram-negative bacteria are known to adhere to components of the extracellular matrix (ECM) and are variable in thickness (from 2 up to 7 nm) and are composed by non-covalent homo-polymerized major subunits.
In contrast, pili in Gram-positive bacteria are long and flexible appendages with a diameter ranging from 3 to 10 nm composed by covalently linked major and ancillary subunits (assembled by specialized sortase enzymes) which are ultimately covalently linked to the peptidoglycan (Telford *et al.*, 2006). Gram-positive pili are usually encoded by islets inserted into specific genomic regions, varying in length and genetic organization, but sharing characteristic features; all pilus encoding islets contain genes coding for: LPxTG or LPxTG-like surface anchored proteins, constituting the pilus major and minor components; and sortase enzymes, which are specialized transpeptidases involved in pilus assembly.
Pilus-like strctures on the surface of Gram-positive bacteria were first detected in *Corynebacterium renale*. More recently, surface appendages were reported to be present in *Actinomyces naeslundii* and were subsequently found in other species, including *Corynebacterium diphtheriae*, and many oral *Streptococcus spp.* (Ton-That *et al.*, 2004b). Finally, pili were also characterized in all three of the principal streptococcal pathogens that cause invasive disease in humans — group A *Streptococcus* (GAS; that is, *Streptococcus pyogenes*), group B *Streptococcus* (GBS; that is, *Streptococcus agalactiae*) and *Streptococcus pneumoniae* — in which they have been shown to have key roles in the adhesion and invasion process and in pathogenesis.
In detail, through the analysis of *Streptococcus pneumoniae* genomes, two genomic elements containing genes typical of Gram-positive pili islets have been identified: pilus islet 1 (PI-1) and pilus islet 2 (PI-2) (Barocchi *et al.*, 2006; Bagnoli *et al.*, 2008) (Fig. 8).

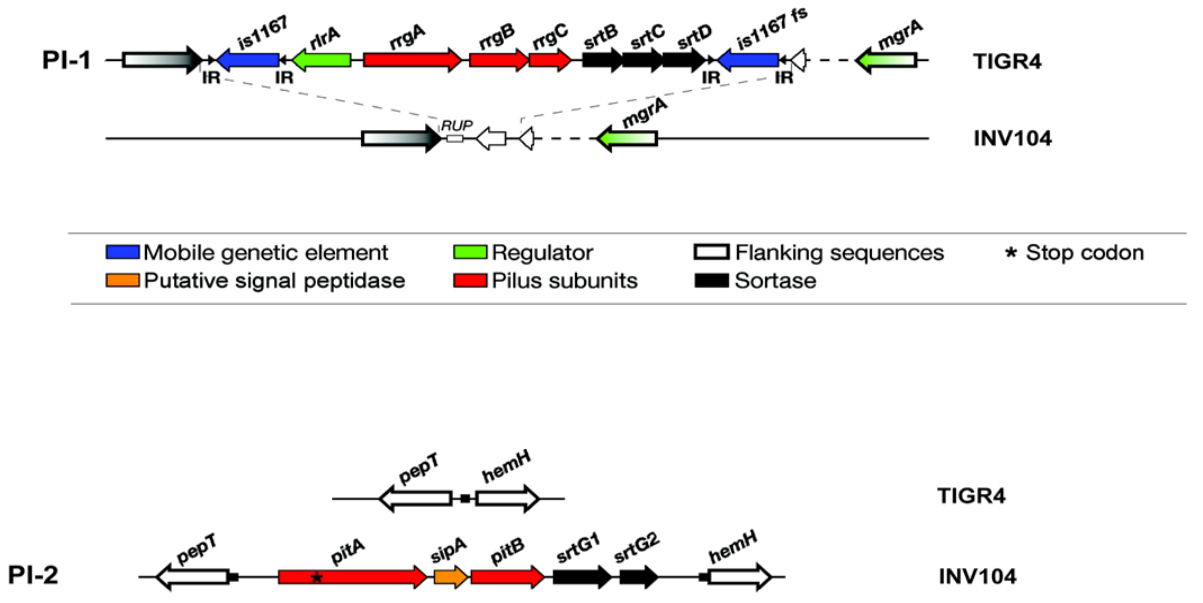


Fig. 8: The genomic organization of pilus-encoding islets in *S. pneumoniae.* Schematic representation of PI-1 and PI-2 (Bagnoli *et al.*, 2008).

Both islets were verified to code for pili on the surface of pneumococcal bacteria (pilus-1 and pilus-2) and to be not widespread in the *S. pneumoniae* population. Interestingly, molecular epidemiology and genomic analyses demonstrated that unlike *S. pneumoniae*, all *S. pyogenes* and *S. agalactiae* isolates encode at least one functional pilus, suggesting different roles for pili in these major streptococcal pathogens (Falugi *et al.*, 2008; Moschioni *et al.*, 2008).
In detail, PI-1, first identified in the genome of the TIGR4 strain, is a chromosomal region of approximately 12 kb, flanked by two insertion sequences (IS1167) and comprising seven genes: *rlrA*, encoding a RofA-like transcriptional regulator, *rrgA*, *rrgB* and *rrgC*, which encode the LPxTG cell-wall anchored pilin subunits, and *srtC-1*, *srtC-2* and *srtC-3* coding for the three sortases involved in the linkage and assembly of the pilus structure on the bacterial surface.
It has been demonstrated that the major pilin subunit, forming the backbone of the pilus, is RrgB, while RrgA and RrgC are the two ancillary pilins, probably localized at the opposite extremes of the pilus; RrgA is the pilus adhesin, while the role of RrgC is still under investigation (Barocchi *et al.*, 2006; Hilleringmann *et al.*, 2008) (Fig. 9).
The deletion of all three pilus-associated sortase genes, *srtC-1*, *srtC-2* and *srtC-3*, completely prevent pilus biogenesis, and expression of SrtC-1 alone is sufficient to covalently associate RrgB subunits to one another as well as linking the RrgA adhesion and RrgC to the polymer (Falker *et al.*, 2008; LeMieux *et al.*, 2008).
In particular, Falker's studies suggested that, SrtC-1 and SrtC-2 act as redundant pilus subunit polymerases, with SrtC-1 processing all three pilus subunits proteins, while SrtC-2 only RrgA and RrgB. In contrast, SrtC-3 seems to have no pilus polymerase activity, but appears to be required for wild type focal presentation of the pili on the bacterial surface.

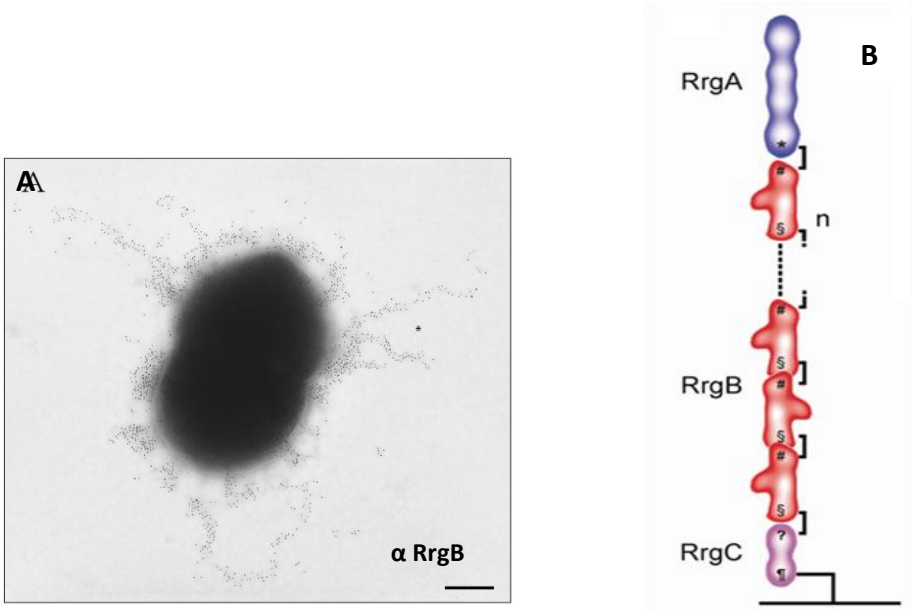


Fig. 9: Detection of pneumococcal surface exposed pilus structures. A) Immunogold localization of RrgB in pilus-1 of *S. pneumoniae* TIGR4 using α-RrgB antibodies. B) Schematic *representation* of pilus-1 subunits arrangement.

PI-2 is similar in its global organization and in its sequence homology to FCT-3 pilus islet of *S. pyogenes* (Teldford *et al.*, 2006). PI-2 islet contains five genes coding for two putative sortases (*srtG1* and *srtG2*), a signal peptidase related protein (*sipA*) essential for pilus assembly (but whose role has not yet been established), and two LPxTG-type surface anchored proteins (*pitB* and *pitA*), where PitB is the backbone subunit, and PitA is the ancillary protein. Interestingly, probably due to the different functions of the two pili (studies on the possible role of PI-2 in pathogenesis are currently ongoing); several copies of pilus-1 are present on the surface of the bacteria, whereas only a single copy of pilus-2 is detectable for each diplococcus (Fig. 10).
Independent studies analyzing global *S. pneumoniae* collections demonstrated that in both invasive, acute otitis media and nasopharyngeal clinical isolates the presence of PI-1 was correlated with the genotype of the isolate, as defined by multi locus sequence typing (MLST), but not with the serotype(Fig.11) (Moschioni *et al.*, 2008; Ogunniyi *et al.*, 2007).
Interestingly, pneumococcal pilus 1 has been reported to be a virulence determinant in *in vivo* animal model studies and its components, RrgA, RrgB and RrgC (in particular RrgB) to be highly immunogenic and to elicit different degrees of protection from lethal challenge both in active and passive immunization studies in different mouse model of infection (intra-nasal, intra-peritoneal and intravenous challenge). For this reason, despite their sequence variability and the limited coverage of these antigens, these proteins are currently evaluated as potential candidates to be included in a new-generation protein based pneumococcal vaccine (Gianfaldoni *et al.*, 2007).

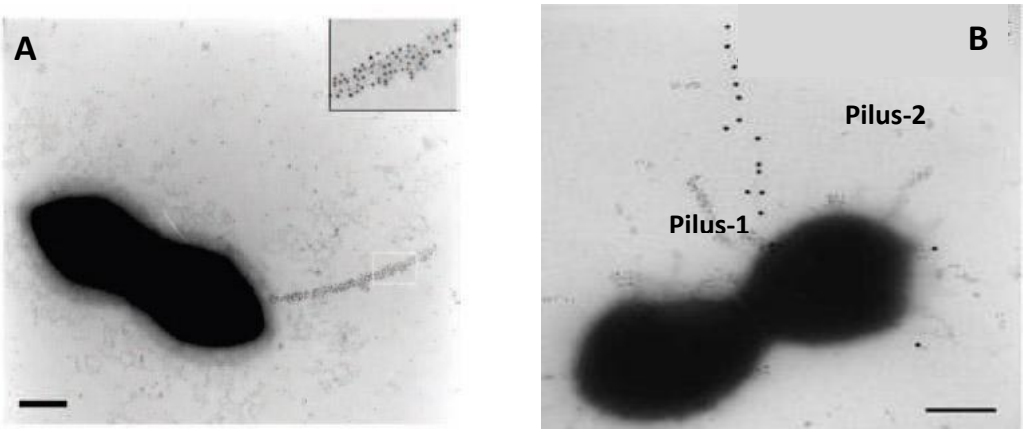


Fig. 10 : Electron-microscopy pictures of pilus-2 structure on the *S. pneumoniae* surface.
A) Immunogold Electron Microscopy localization of PitB on pili of *S. pneumoniae* PN110 whole cells. B) Double immunogold labeling of pilus-1 and pilus-2, respectively, performed with α-RrgB and α-PitB antibodies on 19F Taiwan 14 strain.

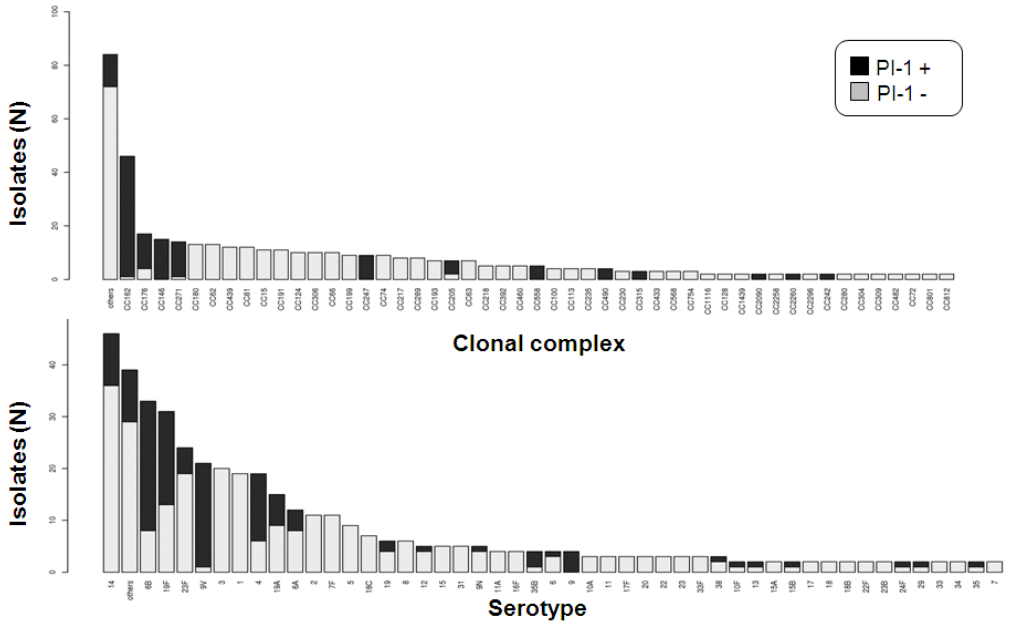


Figure 11. **Distribution of PI-1 in a global *S. pneumoniae* collection.** Histograms showing the relative frequency of PI-1 positive (black) and PI-1 negative (white) strains within a global pneumococcal collection, stratified by Clonal complex (upper histogram) or Serotype. Total number of isolates N=424 (Moschioni et al. 2008).

References

Adam D. (2009). *Issues in pneumococcal disease and pneumococcal conjugate vaccines: highlights of the 27th meeting of ESPID. Brussels, Belgium June, 9-13, 2009.* Vaccine 27(50), 7133-7137

Arrecubieta C., Lopez R., Garcia E. (1994). *Molecular characterization of cap3A, a gene from the operon required for the synthesis of the capsule of Streptococcus pneumoniae type 3: Sequencing of mutations responsible for the unencapsulated phenotype and localization of the capsular cluster on the pneumococcal chromosome*. J. Bacteriol. 176, 6375–6383.

Bagnoli F., Moschioni M., Donati C., Dimitrovska V., Ferlenghi I., Facciotti C. *Et al.* (2008). *A second pilus type in Streptococcus pneumoniae is prevalent in emerging serotypes and mediates adhesion to host cells.* J. Bacteriol. 190, 5480-5492

Barocchi M.A., Ries J., Zogaj X., Hemsley C., Albiger B., Kanth A. *et al.* (2006). *A pneumococcal pilus influences virulence and host inflammatory responses.* Proc. Natl. Acad. Sci. USA 103, 2857-2862

Bentley S.D., Aanensen D.M., Mavroidi A., Saunders D., Rabbinowitsch E., Collins M. *et al.* (2006). *Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes.* PloS Genet. 2, e31

Bogaert D., van Belkum A., Sluijter M., Luijendijk A., de Groot R., Rümke H.C., Verbrugh H.A., Hermans P.W. (2004). *Colonization by Streptococcus pneumoniae and Staphilococcus auresu in healthy children*. Lancet 363 (9424), 1871-1872

Cartee R.T., Forsee W.T., Jensen J.W., Yother J. (2001). *Expression of the Streptococcus pneumoniae type 3 synthase in Escherichia coli. Assembly of type 3 polysaccharide on a lipid primer*. J. Biol. Chem. 276, 48831–48839.

Chan M.S., Maiden M.C., and Spratt B.G. (2001). *Database-driven multi locus sequence typing (MLST) of bacterial pathogens*. Bioinformatics 17 (11), 1077-1083

Crimmins E.M., and Finch C.E. (2006). *Infection, inflammation, height and longevity*. Proc. Natl. Acad. Sci. U.S.A. 103, 498-503

Dillard J.P., Vandersea M.W., Yother J. (1995). *Characterization of the cassette containing genes for type 3 capsular polysaccharide biosynthesis in Streptococcus pneumoniae*. J. Exp. Med. 181, 973–983

Falker S., Nelson A.L., Morfeldt E., Jonas K., Hultenby K., Ries J. *et al.* (2008). *Sortase-mediated assembly and surface topology of adhesive pneumococcal pili.* Mol. Microbiol. 70, 595-607

Falugi F., Zingaretti C., Pinto V., Mariani M., Grandi G. *et al.* (2008). *Sequence variation in group A Streptococcus pili and association of pilus backbone types with lancefield T serotypes.* J. Infect. Dis. 198, 1834-1841

Garcia E., Llull D., Munoz R., Mollerach M., Lopez R. (2000). *Current trends in capsular polysaccharide biosynthesis of Streptococcus pneumoniae*. Res. Microbiol. 151, 429–435

Gianfaldoni C., Censini S., Hilleringmann M., Moschioni M., Facciotti C., Pansegrau W. *et al.* (2007). *Streptococcus pneumoniae pilus subunits protect mice against lethal challenge.* Infect. Immun. 75, 1059-1062

Givon-Lavi N., Fraser D., Porat N., and Dagan R. (2002). Spread of Streptococcus pneumoniae and antibiotic-resistant S. pneumoniae from day-care center attendees to their younger siblings: J. Infect. Dis. 186 (11), 1608-1614

Hausedorff W.P., Bryant J., Kloek C., Paradiso P.R., and Siber G. R. (2000). *The contribution of specific pneumococcal serogroups to different disease manifestations: implications for conjugate vaccine formulation and use,* part II. Clin. Infect. Dis. 30, 122-140

Higginson J. (1977). *The role of the pathologist in environmental medicine and public health.* Am. J. Pathol. 86, 460-484

Hilleringmann M., giusti F., baudner B.C., Masignani V., Covacci A., Rappuoli R. *et al.* (2008). *Pneumococcal pili are composed of profilaments exposing adhesive clusters of RrgA.* PLoS Pathog. 4, e1000026

Jones K.E., Patel N.G., Levy M.A., Storeygard A., Balk D., Gittleman J.L., and Daszak P. (2008). *Global trends in emerging infectious diseases*. Nature 451, 990-993

Kolkman M.A., van der Zeijst B.A., Nuijten P.J. (1998). *Diversity of capsular polysaccharide synthesis gene clusters in Streptococcus pneumoniae*. J. Biochem. (Tokyo) 123, 937–945
Last JM: A dictionary of Epidemiology, ed. 2 New York, Oxford University Press, 1988

LeMieux J., Woody S., and Camilli A. (2008). *Roles of the sortases of Streptococcus pneumoniae in assembly of the RlrA pilus.* J. Bacteriol. 190, 6002-6013

Levin B.R., Lipsitch M., and Bonhoeffer S. (1999). *Population biology, evolution and infectious diseases: convergence and synthesis.* Science 283, 806-809

Llull D., Munoz R., Lopez R., Garcia E. (1999). *A single gene (tts) located outside the cap locus directs the formation of Streptococcus pneumoniae type 37 capsular polysaccharide. Type 37 pneumococci are natural, genetically binary strains*. J. Exp. Med. 190, 241–251

Mausner J.S., and Bahn A.K. (1974). *Epidemiology: an Introductory Text,* p. 3-4. The W.B. Saunders Co, Philadelphia, Pa.

Mavroidi A., Aanensen D.M., Godoy D., Skovsted I.C., Kaltoft M.S., Reeves P.R. *et al.* (2007a). *Genetic relatedness of the Streptococcus pneumoniae capsular biosynthetic loci.* J. Bacteriol. 189, 7841-7855

Moschioni M., Donati C., Muzzi A., Masignani V., Censini S., Hanage W.P. *et al.* (2008). *Streptococcus pneumoniae contains 3 rlrA pilus variants that are clonally related.* J. Infect. Dis. 197, 888-896

Navarre W.W., and Schneewind O. (1999). *Surface proteins of gram-negative bacteria and mechanisms of their targeting to the cell wall envelope.* Microbiol. Mol. Biol. Rev. 63 (1), 174-229

Ogunniyi A.D., Grabowicz M., Briles D.E., Cook J., and Paton J.C. (2007). *Development of a vaccine against invasive pneumococcal disease based on combinations of virulence proteins of Streptococcus pneumoniae.* Infect. Immun. 75, 350-357

Park I.H., Pritchard D.G., Cartee R., Brandao A., Brandileone M.C. and Nahm M.H. (2007). *Discovery of a new capsular serotype (6C) within serogroup 6 of Streptococcus pneumoniae*. J. Clin. Microbiol. 45, 1225-1233

Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico B., Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B. *et al.* (2000). *Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing.* Science 287, 1816-1820

Proft T., and Baker E.N. (2009). *Pili in Gram-negative and Gram-positive bacteria – structure, assembly and their role in disease.* Cell. Mol. Life Sci. 66, 613-635

Rappuoli R. (2004). *From Pasteur to genomics: progress and challenges in infectious diseases*. Nat. Med. 10, 1177-1185

Rappuoli R. (2007). *Bridging the knowledge gaps in vaccine design*. Nat. Med. 10, 1187-1185

Rappuoli R., Miller H.I., and Falkow S. (2002). Medicine. *The intangible value of vaccination*. Science 297, 937-939

Robbins J.B., Schneerson R., Trollfors B., Sato H., Sato Y., Rappuoli R., and Keith J.M. (2005). *The diphtheria and pertussis components of diphtheria-tetanus toxoids-pertussis vaccine should be genetically inactivated mutant toxins*. J.Infect. Dis. 191, 81-88

Rose M., and Zielen S. (2009). *Impact of infant immunization programs with pneumococcal conjugate vaccine in Europe*. Expert Rev Vaccines 8(10), 1351-1364

Ryan M.W., and Antonelli P.J. (2000). *Pneumococcal antibiotic resistance and rates of meningitis in children.* Laryngoscope 110 (6), 961-964

Scarselli M., Giuliani M.M., Adu-Bobie J., Pizza M., and Rappuoli R. (2005). *The impact of genomics on vaccine design*. Trends in Biotech. 23(2), 84-91

Sørensen U.B., Henrichsen J., Chen H.C., Szu S.C. (1990). *Covalent linkage between the capsular polysaccharide and the cell wall peptidoglycan of Streptococcus pneumoniae revealed by immunochemical methods*. Microb. Pathog. 8, 325–334.

Sørensen U.B.S., Agger R., Bennedsen J., and Henrichsen J. (1984). *Phosphorylcholine determinants in six pneumococcal capsular polysaccharides detected by monoclonal antibody.* Infect. Immyn. 43, 876-878

Spratt B.G., Hanage W.P., Li B., Aanensen D.M., and Feil E.J. (2004). *Displaying the relatedness among isolates of bacterial species – the eBURST approach*. FEMS Microbiol. Lett. 241, 129-134

Telford J.L., Barocchi M.A., Margarit I., Rappuoli R., and Grandi G. (2006). *Pili in gram-positive pathogens.* Nat. Rev. Microbiol. 4, 509-519
Thon-That H., Marraffini L. A., and Schneewind O. (2004). *Sortases and pilin elements involved in pilus assembly of Corynebacterium diphtheriae.* Mol. Microbiol. 53, 251-261

Ton-That H., Marraffini L.A., and Schneewind O. (2004a). *Protein sorting to the cell wall envelope of Gram-positive bacteria.* Biochim. Biophys. Acta 1694, 269-278

Ton-That H., Marraffini L.A., and Schneewind O. (2004b). *Sortases and pilin elements involved in pilus assembly of Corynebacterium diphtheriae.* Mol. Microbiol. 53, 251-261

Van Dam J.E.G., Fleer A., and Snippe H. (1990). *Immunogenicity and immunochemistry of Streptococcus pneumoniae capsular polisaccharides.* Antonie Leeuwenhock 58, 1-47

Waite R.D., Penfold D.W., Struthers J.K., Dowson C.G. (2003). *Spontaneous sequence duplications within capsule genes cap8E and tts control phase variation in Streptococcus pneumoniae serotypes 8 and 37*. Microbiology 149, 497–504

Williams B.G., Gouws E., Boschi-Pinto C., Bryce J., and Dye C. (2002). *Estimates of world-wide distribution of child deaths from acute respiratory infections.* Lancet Infect. Dis. 2, 25-32

World Health Organization (2007). *Pneumococcal conjugate vaccine for childhood immunization – WHO position paper.* Wkly. Epidemiol. Record 82, 93-104

Yother J. Capsule. In: Tuomanen EI, editor. *The pneumococcus.* Washington (D. C.): ASM Press; 2004

Chapter 1

Assessment of Streptococcus pneumoniae PI-1 prevalence in carried and transmitted isolates from mother-infant pairs on the Thailand-Burma border.*

*Based on

Paul Turner[1,2,3], Sara Melchiorre[4], Monica Moschioni[4], Michèle A. Barocchi[4], Claudia Turner[1,2,3], Wanitda Watthanaworawit[1,2], Napaphat Kaewcharernnet[1,2], François Nosten[1,2,3], and David Goldblatt[5]

[1]Shoklo Malaria Research Unit, Mae Sot 63110, Thailand
[2]Mahidol-Oxford Tropical Medicine Research Unit, Bangkok 10400, Thailand
[3]Centre for Tropical Medicine, University of Oxford, Oxford OX3 7LJ, United Kingdom
[4]Novartis Vaccines and Diagnostics s.r.l., Siena 53100, Italy
[5] Immunobiology Unit, Institute of Child Health, University of London, London WC1N 1EH, United Kingdom

Introduction

Streptococcus pneumoniae is a significant global pathogen (O'Brien *et al.*, 2009). Additionally, pneumococcus is a nasopharyngeal commensal and infants are frequently colonised (Hill *et al.*, 2008; Bogaert *et al.*, 2004). Various cell-surface components, including pneumococcal surface adhesin A (PsaA), choline-binding protein A (CbpA), pneumococcal serine-rich repeat protein (PsrP) and pneumococcal adherence-virulence factor A (PavA) have been demonstrated to contribute to nasopharyngeal adherence and colonisation, but their mechanisms of action remain incompletely understood (Rajam *et al.*, 2008; Rosenow *et al.*, 1997; Shivshankar *et al.*, 2009; Pracht *et al.*, 2005). The RrgA subunit of the surface exposed pilus-1 filamentous structure also enhances pneumococcal adherence to respiratory epithelial cells *in vitro* (Nelson *et al.*, 2007; Barocchi *et al.*, 2006). Pneumococcal pilus-1 is encoded by the pilus islet-1 (PI-1; *rlrA* islet) and is composed of three subunits, RrgA, RrgB, and RrgC (Barocchi *et al.*, 2006; Hava *et al.*, 2002; LeMieux *et al.*, 2006). Since immunization with pilus antigens is protective against lethal intraperitoneal challenge in a mouse model, pilus subunits are regarded as potential candidates for inclusion in a protein-based pneumococcal vaccine (Gianfaldoni *et al.*, 2007). However, PI-1 presence is not universal: studies of predominantly invasive pneumococci have found that PI-1 is present in isolates from a limited number of serotypes (particularly those included in the 7-valent conjugate vaccine, PCV7) and that its presence correlates with genotype by Multi Locus Sequence Typing (MLST) (Basset *et al.*, 2007; Moschioni *et al.*, 2008; Aguiar *et al.*, 2008; Moschioni *et al.*, 2010; Imai *et al.*, 2010; Donati *et al.*, 2010). Given the prevalence of asymptomatic colonisation, and that serotype/genotype structure of invasive pneumococci is not stable (Black *et al.*, 2010), further work to describe the prevalence and function of the pilus-1 in pneumococcal carriage strains is warranted.
With this respect, our aims were to: i) determine PI-1 prevalence in pneumococcal carriage isolates from infants and their mothers; ii) determine which PI-1 positive clones were associated with carriage in the study region; iii) explore the possible functions of the pilus *in vivo*, and in particular evaluate whether PI-1 presence had an effect on pneumococcal transmission or carriage duration.

Material and Methods

Pneumococcal isolates

Isolates characterized in this work were collected from 2007-2010, during a longitudinal carriage and pneumonia study of 965 infants at Maela refugee camp, Thailand. Maela was established in 1984, and ~40,000 Burmese refugees, predominantly of Karen ethnicity, live in a 4km^2 area. Pneumococcal vaccines are not available in the camp. Nested within the study was a sub-cohort of 234 infant-mother pairs who were studied in more detail: between 1 - 24 months of age, these infants had a nasopharyngeal swab (NPS) and blood taken at monthly surveillance visits and their mothers had an NPS taken at the same time. NPSs were processed using the WHO pneumococcal colonization protocol as previously described (O'Brien *et al.*, 2003; Turner *et al.*, 2011). All morphologically distinct pneumococcal isolates from

each swab culture were serotyped by latex agglutination (Adegbola *et al.*, 2006). Non-typeable (NT) pneumococci, either morphologically typical pneumococcal colonies (likely encapsulated organisms) or rough colonies (likely non-encapsulated organisms) that were non-reactive with typing antisera, were confirmed by bile solubility and absent capsular swelling with Omniserum (Statens Serum Institute, Denmark). Serotype 6C was identified by PCR (Park *et al.*, 2007).

Mother and infant pneumococcal carriage patterns for the first twelve months of follow-up were reviewed and isolates chosen for further analysis using the definitions described below.

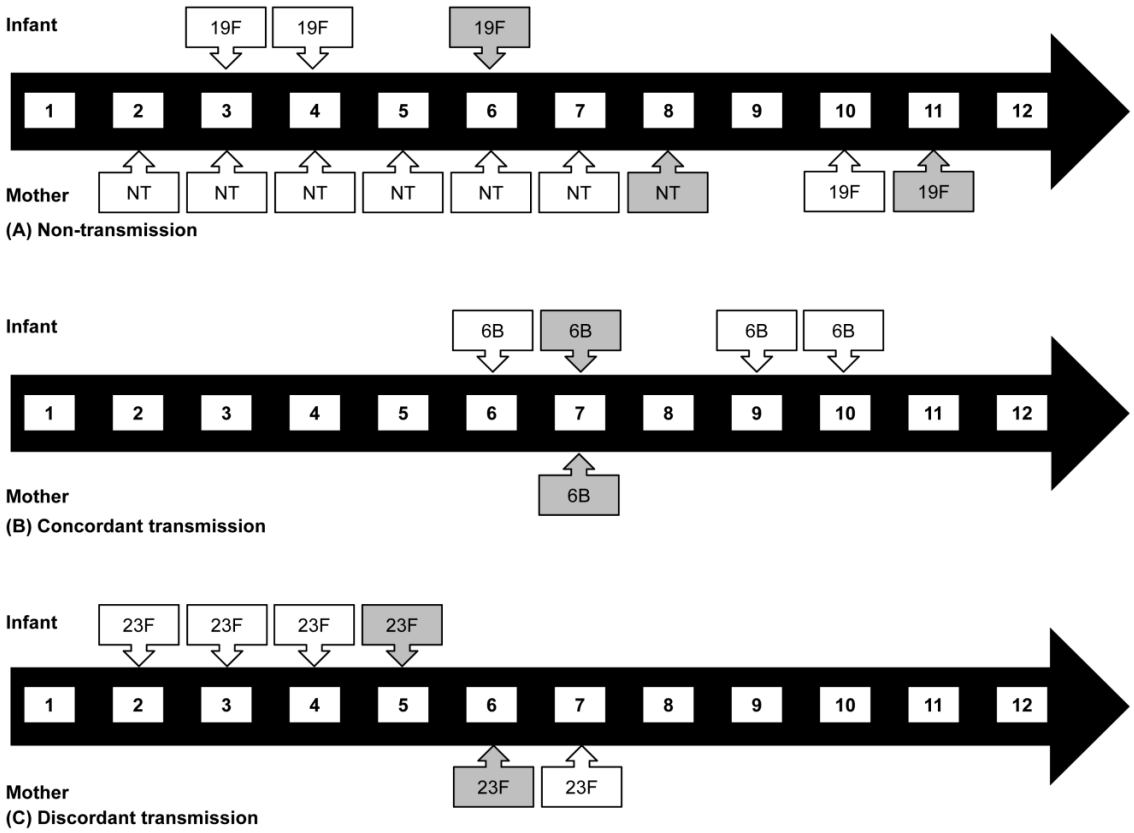
Definition of carriage and transmission

We defined a carriage episode within a mother-infant pair as the identification of the same pneumococcal serotype in ≥2 NPS (from mother and/or infant), separated by ≤2 NPS negative for that serotype. This definition was relaxed for serotypes 1 and 5, known to be carried for very short durations, where a single positive NPS could define a carriage episode.

Transmission of a pneumococcus was defined as the presence of an identical pneumococcal serotype and MLST genotype in both mother and infant at the same visit ("concordant transmission") or when isolated from the mother and the infant during a carriage episode but not at the same visit ("discordant transmission"). A non-transmitted pneumococcus was a carried serotype that never appeared in the other member of the mother-infant pair during the carriage episode as defined above (Table 1 and Supplementary Fig. 1).

Carriage episode. The same serotype identified in at least two NPS, separated by ≤2 negative NPS (exception: serotypes 1 and 5).
Non-transmitted. A carried pneumococcus that never appears in the other member of the mother-infant pair.
Transmitted – concordant. Identical pneumococci (serotype + MLST) identified from a mother and baby at the same visit.
Transmitted – discordant. Identical pneumococci identified from a mother and baby during a carriage episode (but not on the same visit).

Table 1: Definition of carriage and transmission.



Supplementary Fig.1: Example of pneumococcal transmission event. (A) non-transmission of serotypes 19F and NT; (B) concordant transmission of serotype 6B; (C) discordant transmission of serotype 23F. Horizontal arrows represent an infant's first year of life, with the numbers indicating age in months. Pneumococcal serotypes shown above these arrows represent carried isolates from the infant and those below the arrows are from the mother. Shading highlights the criteria used to select isolates for PI-1 PCR and MLST.

PI-1 detection

PI-1 presence was determined as previously described (Moschioni *et al.*, 2008). Briefly, PCRs were performed directly from bacteria using the primers listed in the Table 2. Primers were designed on conserved regions on the boundaries of PI-1 (459for, 470rev) and within PI-1 (P01rev, P11for, P08for, P08rev).

Forward primer	Sequence (5'→3')	Reverse primer	Sequence (5'→3')
459 for	AACTGAATTGACACAACGTGTCTT	470 rev	GCCACACAAGATGTTGATGCTTTT
459 for	AACTGAATTGACACAACGTGTCTT	P01 rev	AGCGACAAGCCACTGTATCATATT
P08 for	TGAGATTTTCTCGTTTCTCTTAGC	P08 rev	AATAGACGATGGGTATTGATCATGT
P11for	GCCATTTGGATCAGCTAAAAGTT		

Table 2: PI-1 PCR primer sets

Multi Locus Sequence Typing

MLST was performed as previously described (Enright *et al.*, 1998). Briefly, PCR amplifications were performed directly from the bacteria using the standard primer pairs. Sequences were obtained on both strands using an ABI 3730xl DNA Analyzer (Life Technologies Corporation, USA). Sequence type (ST) was determined using the MLST website (http://spneumoniae.mlst.net). eBURST (http://spneumoniae.mlst.net/eburst/) was run with default settings on the entire MLST database and each ST assigned to a clonal complex (CC) (Feil *et al.*, 2004). CCs were named in accordance with the ST number of the founder predicted by eBURST.

Statistical analysis

Statistical analyses were carried out using STATA 10.1 (StataCorp, USA). Chi-squared test, Fisher's exact test, and odds ratios were used to compare proportions. Carriage duration was estimated by survival analysis methods, using the Log-rank test to compare groups. First pneumococcal acquisition was defined as midpoint between the last negative swab and the first positive swab for a given serotype. Termination of the carriage episode was similarly defined as the midpoint between the first of two consecutive negative swabs and the last positive swab for the serotype (Hill *et al.*, 2008).

Study ethics

Ethical approval for the carriage study was granted by Mahidol University, Thailand and Oxford University, UK.

Results

Pneumococcal isolates and PI-1 prevalence

In the first twelve months of follow-up, 4.927 surveillance NPS were collected from the 234 mother-infant pairs (85% of expected). 2,496 pneumococci were cultured and 896 (35.9%) of these isolates were included in the current work. Thirty-four serotypes were contained in the strain selection: these serotypes accounted for 90.1% (2.250/2.496) of isolates in the entire cohort strain collection.
We successfully determined the PI-1 status of 887/896 (98.9%) isolates: 35.2% (312/887) were positive. PI-1 positive isolates were restricted to 10 serotypes: 4, 6A, 6B, 9, 14, 19F, 19A, 23F, 33C, and NT (Fig. 1).

PI-1 and transmission

To analyse whether PI-1 presence had an effect on pneumococcal transmission within the mother-infant pairs, a panel of strains was selected and analysed by both PI-1 PCR and MLST (to define genotype). We elected to study serotypes contained

in, or related to, current conjugate vaccines (PCV13 + 6C) as well as non-typeable pneumococci which were overrepresented in our population (the commonest pneumococcal “type” identified from mother swabs and the third commonest in infants). In the first year of follow-up, these serotypes accounted for 69.6% and 59.8% of pneumococci carried by infants and mothers, respectively. In the case of a potentially transmitted pneumococcal serotype, i.e. isolates carried by both mother and infant at the same visit, or at
sequential visits, single isolates from both mother and infant were selected for study. For non-transmitted pneumococci, i.e. pneumococci carried by only the mother or infant, a single isolate from the carriage episode was studied. Additionally, we analyzed every isolate from carriage episodes of all serotypes in eight mother-infant pairs to determine the clonality of pneumococcal carriage.

Of the 489 isolates selected for the primary analysis, PI-1 PCR was uninterpretable in six, resulting in 483 analyzable pneumococcal isolates. Overall, 219/483 (45.3%) of these isolates were PI-1 positive. PI-1 was found in non-typeable pneumococci and in 6/14 serotypes studied, all of which, apart from one 19A isolate, were PCV7 serotypes (Table 3).

PI-1 was found most frequently in 19F (91%) and 23F (77%) although these serotypes, according to the MLST and eBURST analyses, were dominated by a single clonal complex each (CC271 – 19F; CC802 – 23F). Overall, PI-1 positive pneumococci were members of a restricted group of clonal complexes (Figure 2).
By serotyping a representative isolate alone, 253/483 (52.4%) pneumococci would have been classified as transmitted (153 concordant; 100 discordant) and 230/483 (47.6%) as non-transmitted. Combining MLST and serotype, nine concordant and eight discordant “transmitted” isolates (17/483; 3.5%) were reclassified as non-transmitted, resulting in a final total of 236/483 (48.9%) transmitted and 247/483 (51.1%) non-transmitted pneumococci.

PI-1 presence was not correlated with transmission: 47.5% (112/236) of transmitted and 43.3% (107/247) of non-transmitted pneumococci were PI-1 positive, with an odds ratio of 1.2 (95% CI: 0.8 – 1.7; p=0.4), and no significant differences at the individual serotype level.

We analysed an additional 91 pneumococcal isolates, from seven mother-infant pairs and one mother-twin infant unit, to confirm that sequential isolates of the same serotype in an individual, and isolates of the same serotype in a mother-infant pair, were indeed identical. In all mother-infant pairs each serotype carried was represented by a single ST per carriage/transmission episode with three exceptions (Supplementary Table 2).

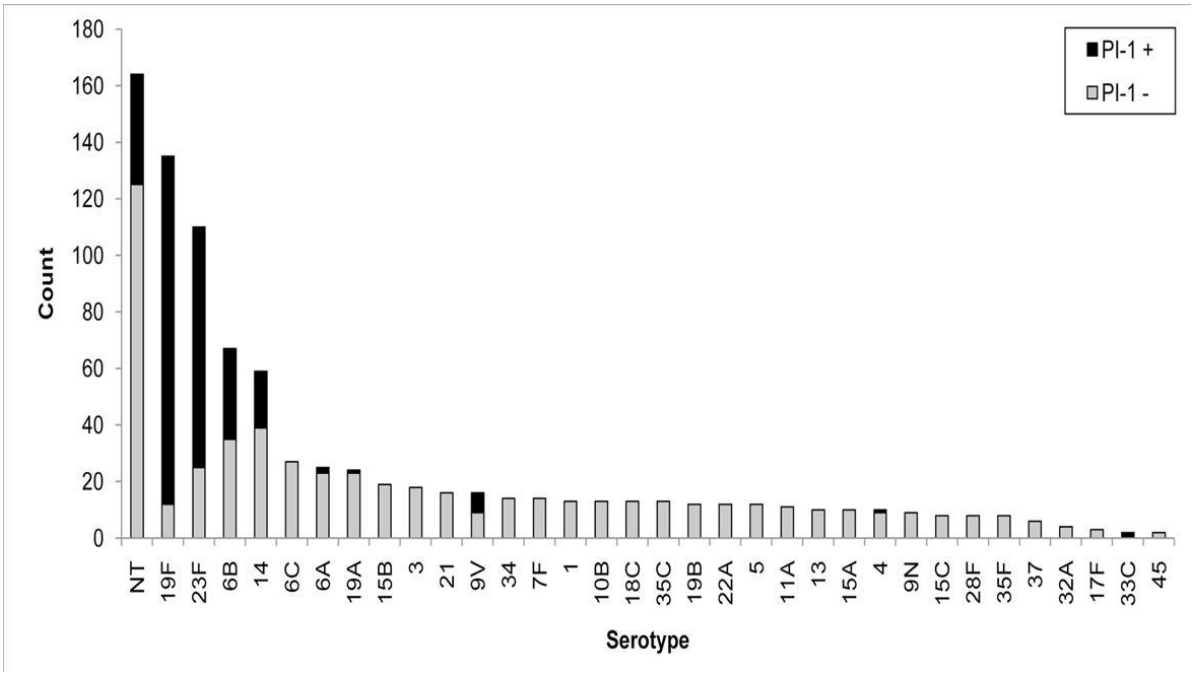


Fig. 1: PI-1 presence by serotype in 887 carried pneumococci (includes all commonly carried serotypes).

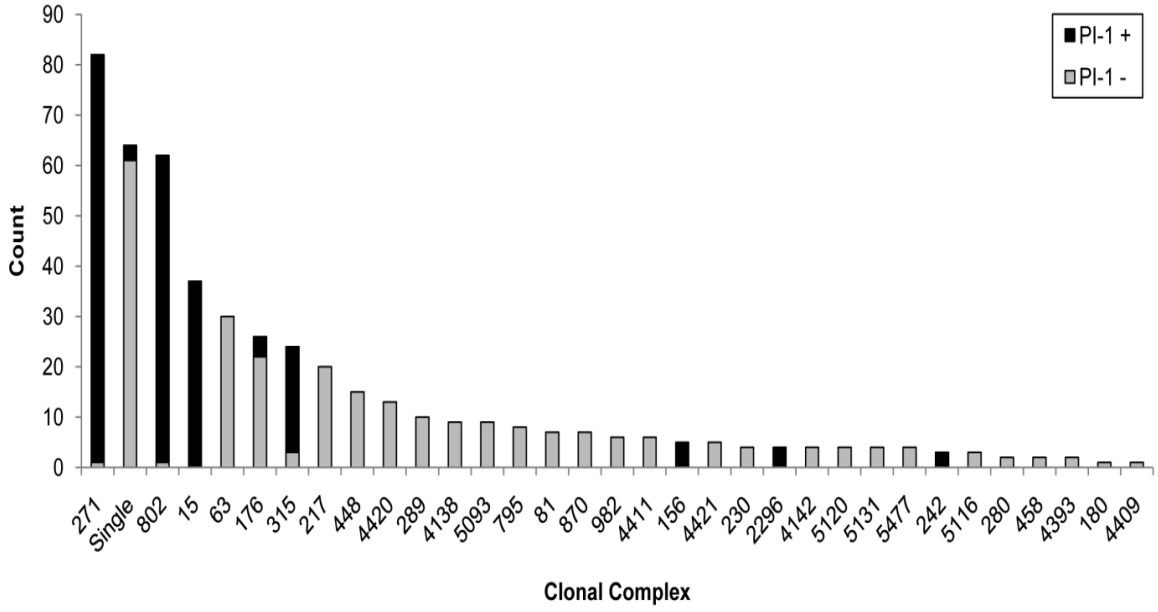


Fig. 2. PI-1 presence by clonal complex in 483 carried pneumococci (PCV13 serotypes plus 6C and NT only).

Serotype	Total, N	Transmitted*, N (% within serotype)	PI-1 present, N (% within serotype)
1	12	4 (33.3)	0 (0.00)
3	11	4 (36.4)	0 (0.00)
4	4	2 (50.0)	0 (0.00)
5	10	2 (20.0)	0 (0.00)
6A	16	6 (37.5)	0 (0.00)
6B	54	27 (50.0)	29 (53.7)
6C	15	4 (26.7)	0 (0.00)
7F	5	2 (40.0)	0 (0.00)
9V	8	4 (50.0)	5 (62.5)
14	46	28 (60.9)	15 (32.6)
18C	7	4 (57.1)	0 (0.00)
19F	89	50 (56.2)	81 (91.0)
19A	14	4 (28.6)	1 (7.1)
23F	82	38 (46.3)	63 (76.8)
NT	110	57 (51.8)	25 (22.7)
Total	483	236 (48.9)	219 (45.3)

Table 3. Pneumococcal serotype distribution and isolate transmission category in mother-infant pairs (PCV13 serotypes plus 6C and NT).

PI-1 and first pneumococcal carriage episode duration in infants

To analyse the correlation between PI-1 presence and the duration of infant carriage, we determined the PI-1 status of 316 isolates from first carriage episodes of all common serotypes (i.e. all first episodes of carriage of each serotype, not necessarily an infant's first ever pneumococcal carriage episode). We inferred the PI-1 status for the entire carriage episode from the PI-1 PCR result of a single isolate. We included all carriage episodes, even where a pneumococcal serotype was identified only in a single NPS.

Given the well described association of shorter carriage in individuals with prior pneumococcal exposure, we subsequently focused our analysis on only first ever episodes of carriage (i.e. one per infant) (Hill *et al.*, 2008). From the 316 carriage episodes selected, we identified 216 analysable first ever carriage episodes: PI-1 PCR data was available for at least one isolate in 180 of these (90 PCV13 serotypes, 58 non-vaccine serotypes, and 32 non-typeable pneumococci). Median and mean durations of the first pneumococcal carriage episode were 63 days (95% CI: 61 – 91) and 105 days (95% CI: 91 – 119) respectively, but varied considerably by serotype (Table 4).

Comparing the 180 carriage episodes where PI-1 PCR results were available, PI-1 positive carriage episodes (51/180; 28.3%) were significantly longer than those associated with a PI-1 negative organism (median 152 days (95% CI: 93 – 213) vs. 61 days (95% CI: 57 – 90); mean 177 days (95% CI: 137 – 218) vs 84 days (95% CI: 72 – 97); $p<0.0001$). However, the analysis of carriage duration was confounded by serotype. In our study serotypes 19F and 23F had the longest duration of carriage and were both predominantly PI-1 positive. To analyze the relative contribution of serotype and PI-1 to carriage duration we fitted a Cox regression model and found that, stratifying carriage by serotype, PI-1 presence was not associated with a significant change in carriage duration (hazard ratio 0.7 (95% CI: 0.4 – 1.2; p=0.2)).

Serotype	Number of episodes	Median carriage duration, days (95% CI)	Mean carriage duration, days (95% CI)
6B	13	121 (90 – 153)	119 (93 – 145)
14	8	62 (30 – 151)	86 (45 – 128)
19F	21	213 (63 – 243)	231 (154 – 308)
23F	21	184 (62 – 277)	176 (124 – 229)
35F	9	121 (30 – 180)	124 (66 – 182)
NT	33	31 (31 – 61)	72 (44 – 99)

Table 4. First pneumococcal carriage episode duration (six commonest serotypes).

Discussion

This study is the largest investigation of the prevalence and possible function of the PI-1 encoded pilus in carriage isolates of *S. pneumoniae*. It also provides pneumococcal strain data from SE Asia, a densely populated region under-represented in the pneumococcal carriage and disease epidemiology literature. As serotypes are commonly used as the basis for defining transmission, we were interested in the frequency with which a molecular analysis may confound the analysis of transmission. By genotyping, discrepancies were discovered in only 3.5% of transmission events classified by serotype. We also demonstrated that pneumococci of the same serotype are predominantly clonal within a discrete carriage episode. We therefore concluded that serotype can be used to define transmission events and carriage episodes with relatively high confidence, but the inclusion of genotyping is vital to ensure complete accuracy.

Because of the selection criteria for isolates included in the study, we cannot describe the overall pilus prevalence for the population. However, in the 887 carriage isolates analysed the pilus prevalence was 35.2%. Of the 34 serotypes analysed we found that PI-1was only present in non-typeable isolates and nine other serotypes: 65% of PCV7 isolates were PI-1 positive compared with 9% of non-vaccine serotypes. Interestingly, Basset et al., examining nasopharyngeal and invasive pneumococcal isolates from the American Indian collection, also found that PCV7 strains were significantly more likely to be PI-1 positive than non-vaccine serotypes, but the overall proportion of PI-1 positive isolates was slightly lower (Basset *et al.*, 2007). They, along with other authors, demonstrated that PI-1 positive strains were contained within a small number of clonal complexes (Moschioni *et al.*, 2008; Aguiar *et al.*, 2008; Moschioni *et al.*, 2010). Indeed, in the current study we demonstrated that PI-1 positive isolates were clustered predominantly within four dominant clonal complexes (CC15, CC271, CC315, CC802). The strong association between serotype/clonal complex and pilus presence, along with strain selection criteria, and the different regional distribution of the clones may explain the variability in PI-1 prevalence between different studies.

We did not demonstrate a pilus-attributable effect on pneumococcal transmissibility. However, there are limitations to our study which may be important confounders for this analysis. The carriage study was carried out in a densely populated refugee camp where 13% of the population are <5 years old and there is likely to be frequent transmission of nasopharyngeal organisms both within families and between members of the community. Despite using a combination of serotype and sequence type to increase the accuracy of our transmissibility categories, we cannot exclude the possibility that “non-transmitted” strains were effectively transmitted between mother or infant and others but not detected by us. Several studies have documented the clustering of pneumococcal serotypes and genotypes within families, which highlights the difficulty of assigning a definitive “non-transmitted” label to isolates collected from an incomplete household group (Melegraro *et al.*, 2004; Reis *et al.*, 2008). However, we focused on mother-infant transmission since we felt that the absence of a strain in one member of this pair would be the best marker for relative non-transmissibility in the early months of life. In addition, we included both concordant and discordant time-point pneumococcal serotype/ST identifications in the mother-infant pair as “transmitted”, as long as they occurred no more than two months apart. However, the study sampling frequency may have been too low to demonstrate transmission of serotypes carried for very

short durations. Also, in the presence of multiple serotype carriage, a common occurrence in infancy and one that is underestimated by standard culture protocols (Turner *et al.*, 2011), a particular serotype may become undetectable for a period of time before re-emerging as the dominant serotype which may result in incorrect categorisation regarding transmissibility.

Although in the crude analysis PI-1 presence was associated with longer first pneumococcal carriage episode in infants, we could find no significant association between carriage duration and PI-1 at the individual serotype level. This is likely to be the result of the low numbers of carriage episodes of individual serotypes (resulting in wide confidence intervals around the carriage duration estimates) and the restricted number of clones within each serotype. We could have increased the number of carriage episodes included in our analysis by looking at all carriage episodes of each serotype rather than restricting to each infants first ever carriage episodes, but this would have introduced other confounding factors, such as the impact of previous carriage and immune factors in subsequent carriage episode duration.

In conclusion, we found that SE Asian pneumococcal carriage isolates had similar pilus prevalence to previously described strain collections, which is helpful in the on-going assessment of likely global coverage of a pilus subunit-containing vaccine. Despite its known role in pneumococcal attachment, we could not determine a clear impact of PI-1 presence on transmissibility or carriage duration.

Infant age (m)	1	2	3	4	5	6	7	8	9	10	11	12
Pair 1												
Infant			19A (230)	19A (230)	14 (63)	14 (63)	14 (63)	NT (4133)	6C (4420)	6C (4420) NT (4133)	6C (4420)	6C (4420)
Mother			35F (4418)	14 (63)							7F (3545)	7F (3545)
Pair 2												
Infant		NT (4136)	NT (4136)	NT (4136)		NT (4136)	NT (4136)	NT (4136)	NT (4136)	NT (4136)		
Mother	NT (4136)		7F (3545)	7F (3545)	NT (4136)	7F (3545)	7F (3545)	7F (3545) NT (4136)	7F (3545)	NT (4136)	NT (4136)	NT (4136)
Pair 3												
Infant		14 (63)	14 (63)	14 (63)	NT (4451)	14 (63)	14 (63)		19F (81)	22A (910)	19F (81)	22A (910)
Mother	32A (5092)	32A (5092)	32A (5092)	NT (5121)	14 (63)	NT (4451)	14 (4451)	NT (4451)	32A (5092)		NT (4451)	19F (81)
Pair 4												
Infant			19F (4414)	19F (4414)	19F (4414)	19F (4414)	19F (4414)	19F (4414)	23F (802)	19F (4414)	19F (4414)	23F (802)
Mother	37 (447)	37 (447)			37 (447)	37 (447)		37 (447)	37 (447)			19F (4414)
Infant age (m)	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Pair 5 (twins)												
Infant 1	NT (448)	33C (5096)	33C (5096)	17F (5098)	21 (5103)	21 (5103)	19F (5106)	19F (5106)	19F (4414)	19F (4414)	6A (4936)	19F (4414) NT (448)
Infant 2		17F (5098)	21 (5103)	21 (5103)	17F (5098)	21 (5103)	21 (5103)	21 (5103)		19F (4414)	6A (4936)	6A (4936)
Mother	NT (448)		1 (217)			NT (448)						
Pair 6												
Infant		NT (4133)	NT (4133)	19B (5095)	1 (217)	NT (4133)	9V (280)	9V (280)	19F (4414)	19F (4414)	19F (4414)	19F (4414)
Mother	19B (5095)	19B (5095)			NT (4133)			9V (5123)	9V (280)			
Pair 7												
Infant	45 (5097)	45 (5097)	23F (802)	23F (802)	23F (802)		6B (315)		23F (802)	23F (802)	23F (802)	
Mother		23F (802)		23F (802)	23F (802)	6B (315)					23F (802)	
Pair 8												
Infant			23F (4413)			23F (4413)	23F (4413)	23F (4413)	23F (4413)		23F (4413)	
Mother					23F (4413)						23F (4413)	

Supplementary Table 2. Detailed study of pneumococcal carriage in the first year of life in eight mother-infant pairs. Pneumococcal serotypes and their MLST genotype (in parentheses) are given for each nasopharyngeal sampling point.

References

O'Brien KL, Wolfson LJ, Watt JP, Henkle E, Deloria-Knoll M, McCall N, et al. Burden of disease caused by *Streptococcus pneumoniae* in children younger than 5 years: global estimates. Lancet2009;374(9693):893-902.

Hill PC, Cheung YB, Akisanya A, Sankareh K, Lahai G, Greenwood BM, et al. Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian infants: a longitudinal study. Clin Infect Dis2008;46(6):807-14.

Bogaert D, De Groot R, Hermans PW. *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. Lancet Infect Dis2004;4(3):144-54.

Rajam G, Anderton JM, Carlone GM, Sampson JS, Ades EW. Pneumococcal surface adhesin A (PsaA): a review. Crit Rev Microbiol2008;34(3-4):163-73.

Rosenow C, Ryan P, Weiser JN, Johnson S, Fontan P, Ortqvist A, et al. Contribution of novel choline-binding proteins to adherence, colonization and immunogenicity of *Streptococcus pneumoniae*. Mol Microbiol1997 Sep;25(5):819-29.

Shivshankar P, Sanchez C, Rose LF, Orihuela CJ. The *Streptococcus pneumoniae* adhesin PsrP binds to Keratin 10 on lung cells. Mol Microbiol2009 Aug;73(4):663-79.

Pracht D, Elm C, Gerber J, Bergmann S, Rohde M, Seiler M, et al. PavA of *Streptococcus pneumoniae* modulates adherence, invasion, and meningeal inflammation. Infect Immun2005 May;73(5):2680-9.

Nelson AL, Ries J, Bagnoli F, Dahlberg S, Falker S, Rounioja S, et al. RrgA is a pilus-associated adhesin in *Streptococcus pneumoniae*. Mol Microbiol2007;66(2):329-40.

Barocchi MA, Ries J, Zogaj X, Hemsley C, Albiger B, Kanth A, et al. A pneumococcal pilus influences virulence and host inflammatory responses. Proc Natl Acad Sci USA2006;103(8):2857-62.

Hava DL, Camilli A. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. Mol Microbiol2002;45(5):1389-406.

LeMieux J, Hava DL, Basset A, Camilli A. RrgA and RrgB are components of a multisubunit pilus encoded by the *Streptococcus pneumoniae* rlrA pathogenicity islet. Infect Immun2006;74(4):2453-6.

Gianfaldoni C, Censini S, Hilleringmann M, Moschioni M, Facciotti C, Pansegrau W, et al. *Streptococcus pneumoniae* pilus subunits protect mice against lethal challenge. Infect Immun2007;75(2):1059-62.

Basset A, Trzcinski K, Hermos C, O'Brien KL, Reid R, Santosham M, et al. Association of the pneumococcal pilus with certain capsular serotypes but not with increased virulence. J Clin Microbiol2007;45(6):1684-9.

Moschioni M, Donati C, Muzzi A, Masignani V, Censini S, Hanage WP, et al. *Streptococcus pneumoniae* contains 3 rlrA pilus variants that are clonally related. J Infect Dis2008;197(6):888-96.

Aguiar SI, Serrano I, Pinto FR, Melo-Cristino J, Ramirez M. The presence of the pilus locus is a clonal property among pneumococcal invasive isolates. BMC Microbiol2008;8:41.

Moschioni M, De Angelis G, Melchiorre S, Masignani V, Leibovitz E, Barocchi MA, et al. Prevalence of pilus encoding islets among acute otitis media *Streptococcus pneumoniae* isolates from Israel. Clin Microbiol Infect2010;16:1501-7.

Imai S, Ito Y, Ishida T, Hirai T, Ito I, Yoshimura K, et al. Distribution and clonal relationship of cell surface virulence genes among *Streptococcus pneumoniae* isolates in Japan. Clin Microbiol Infect2010.

Donati C, Hiller NL, Tettelin H, Muzzi A, Croucher NJ, Angiuoli SV, et al. Structure and dynamics of the pan-genome of *Streptococcus pneumoniae* and closely related species. Genome Biol2010;11(10):R107.

Black S. The volatile nature of pneumococcal serotype epidemiology: potential for misinterpretation. Pediatr Infect Dis J2010;29(4):301-3.

O'Brien KL, Nohynek H. Report from a WHO working group: standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae*. Pediatr Infect Dis J2003;22(2):133-40.

Turner P, Hinds J, Turner C, Jankhot A, Gould K, Bentley S, et al. Improved detection of nasopharyngeal co-colonization by multiple pneumococcal serotypes using latex agglutination or molecular serotyping by microarray. J Clin Microbiol2011;49(5):1784-9.

Adegbola RA, Hill PC, Secka O, Ikumapayi UN, Lahai G, Greenwood BM, et al. Serotype and antimicrobial susceptibility patterns of isolates of *Streptococcus pneumoniae* causing invasive disease in The Gambia 1996-2003. Trop Med Int Health2006;11(7):1128-35.

Park IH, Park S, Hollingshead SK, Nahm MH. Genetic basis for the new pneumococcal serotype, 6C. Infect Immun2007;75(9):4482-9.

Enright MC, Spratt BG. A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. Microbiology1998;144(11):3049-60.

Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J Bacteriol 2004;186(5):1518-30.

Hill PC, Cheung YB, Akisanya A, Sankareh K, Lahai G, Greenwood BM, et al. Nasopharyngeal carriage of Streptococcus pneumoniae in Gambian infants: a longitudinal study. Clin Infect Dis2008 Mar 15;46(6):807-14.

Melegaro A, Gay NJ, Medley GF. Estimating the transmission parameters of pneumococcal carriage in households. Epidemiol Infect2004;132(3):433-41.

Reis JN, Palma T, Ribeiro GS, Pinheiro RM, Ribeiro CT, Cordeiro SM, et al. Transmission of *Streptococcus pneumoniae* in an urban slum community. J Infect2008;57(3):204-13.

Chapter 2

Point Mutations in wchA of 7F Capsular Locus Are Responsible for Non-Typeability of Two Invasive Streptococcus pneumoniae Isolates.*

*Based on

Sara Melchiorre[1]*, Romina Camilli[2]*, Agostina Pietrantoni[3], Monica Moschioni[1], Francesco Berti[1], Maria Del Grosso[2], Fabiana Superti[3], Michèle Anne Barocchi[1], Annalisa Pantosti[2]

[1]Novartis Vaccines & Diagnostics, Siena,
[2]Department of Infectious, Parasitic and Immunemediated Diseases, Istituto Superiore di Sanità, Rome, Italy
[3]Department of Technology and Health, Istituto Superiore di Sanità, Rome, Italy

Introduction

The Gram-positive pathogen *Streptococcus pneumoniae* is a major cause of community acquired pneumonia as well as of upper respiratory tract infections such as acute otitis media and sinusitis, and invasive diseases like meningitis, bacteremia, and endocarditis. However, *S. pneumoniae* is also a commensal of the upper respiratory tract, especially of young children, which represent the reservoir for pneumococcal transmission within the community.
The ability of the pneumococcus to invade sterile sites is related to the expression of a polysaccharide capsule, which provides a barrier against host-cell mediated phagocytosis and allows bacterial persistence in the blood (Brown *et al.*, 1983). In addition, capsular polysaccharides (CPSs) are immunogenic and antibodies against CPSs provide protection against pneumococcal diseases; therefore, all of the current pneumococcal vaccines are based on a combination of different CPSs, either unconjugated or conjugated to a carrier protein (Gladstone *et al.*, 2011).
To date, 93 capsular types (serotypes) have been described on the basis of their different genetic, biochemical and antigenic properties (Bentley *et al.*, 2006; Calix *et al.*, 2010; Park *et al.*, 2007). Each pneumococcal serotype corresponds to a distinct CPS, synthesized on the bacterial surface by enzymes which are encoded within a single chromosomal locus (*cps*) with few remarkable exceptions (Llull *et al.,* 1999). The sequencing of different *cps* loci revealed the presence of a core of genes common to all capsular types, namely *wzg*, *wzh*, *wzd*, and *wze* (also known as *cpsA, B, C and D*), and of additional genes that are CPS-specific (Bentley *et al.*, 2006).
The different capsular types can be distinguished in the laboratory by reaction with type specific antisera, evidenced by the capsular swelling or Quellung reaction (Austrian R., 1976). Only a minority of *S. pneumoniae* strains does not react with the anticapsular antisera; these strains, defined as non-typeable pneumococci (NTPn), are thought to be non-encapsulated. NTPn are typically isolated from carriage (Andrade *et al.*, 2010; Sa-Leao *et al.*, 2006), but are also found in conjunctivitis (Carvalho *et al.*, 2003), acute otitis media (Xu *et al.*, 2011) and, rarely, in invasive diseases (Beall *et al.*, 2006). Non-typeability can be due to different genetic modifications such as partial or complete loss of the *cps* gene cluster (Hathaway *et al.*, 2004), presence of a novel gene (Baldry *et al.*, 2009), single point mutations (Arrecubieta *et al.*, 1994) or sequence duplications (Waite *et al.*, 2001).
The characterization of NTPn by multi locus sequence typing (MLST), revealed the presence of specific NT pneumococcal lineages which had lost the *cps* (Hanage *et al.*, 2006; Hathaway *et al.*, 2004) as well as of NTPn that were genetically related to encapsulated strains (Andrade *et al.*, 2010; Hathaway *et al.*, 2004).
In this report, we investigated the genetic basis for non-typeability of two NTPn isolates responsible for invasive disease. The strains were isolated in the same Italian hospital during 2006 through the surveillance network of invasive pneumococcal disease (Gherardi *et al.*, 2009) from two cases of bacteremic pneumonia.

Material and Methods

Bacterial strains and growth conditions

S. pneumoniae AP422 and AP426 were isolated from the blood of two adult patients with pneumonia admitted to the same Italian hospital in 2006. *S. pneumoniae* AP425 (clinical isolate, serotype 7F) was used as a control. Strains were grown at 37°C in 5% CO2-enriched atmosphere on Columbia agar plates (Oxoid, Cambridge, UK) with 5% sheep blood or on Tryptic Soy Agar plates (Becton Dickinson, Cockeysville, MA) supplemented with 10 mg/l colistine, 5 mg/l oxolinic acid and 5% sheep blood. For genomic DNA extraction and NMR spectroscopy, liquid cultures were grown statically at 37°C in 5% CO2 in Todd Hewitt Broth (Becton Dickinson, Cockeysville, MA) supplemented with 0.5% (w/w) yeast extract (THYE) until an A_{600} = 0.25.

Serotyping

Serotyping was performed by means of the Latex agglutination test and the Quellung reaction using antisera produced by the Statens Serum Institute (Copenhagen, Denmark).

Multi Locus sequence typing (MLST)

Genomic DNA from pneumococcal strains was prepared with the Wizard Genomic DNA purification kit (Promega, Madison, WI, USA) following the manufacturer's instructions. MLST was performed as previously described (Moschioni *et al.*, 2008).

7F cps detection and sequencing

The amplification and sequencing of 7F *cps* was performed on the genomic DNA of strains AP422, AP426 and AP425 with the primers listed in Table S1. Amplicons were purified with magnetic carboxilate beads (Agentcourt, Bioscience, Beverly, MA, USA); sequencing was carried 6 out by use of an ABI 3730xl DNA Analyzer (Applied Biosystems, Carlsbad, CA, USA). Traces were visually inspected, edited, assembled and analyzed with Vector NTI Advance 10.3.1 (Life technologies, Carlsbad, CA, USA).

Cloning and S. pneumoniae transformation

The shuttle plasmid pMU1328_*Pc_wchA* was obtained as described below. Briefly, the *wchA* gene of the AP425 strain and the constitutive promoter of erythromycin (*Pc*) (Gentile *et al.*, 2011) were amplified with the primers *wchA*_for GTGCGTGGATCCATGGATGAAAAAGGATTGAAAATT, *wchA*_rev CAGCGTGGATCCTCACTTCGCCCCTTCTCTCATAAA, and *Pc*_for GTGCGTGAATTCGAAACAGCAAAGAATGGCGGAAAC, *Pc*_rev CAGCGTGGATCCGTAATCACTCCTTCTTAATTACAA, respectively. The two PCR products, digested with BamHI and EcoRI-BamHI restriction enzymes, respectively, were cloned into pMU1328, containing an erythromycin resistance marker (Achen *et al.*, 1986). The ligations mixtures were then transformed into competent *Escherichia*

coli DH10B and transformants were selected on plates supplemented with erythromycin (100 µg/ml). *Pc_wchA* insertion was confirmed by sequencing.
pMU1328_*Pc_wchA* was then transformed in AP422 and AP426 by using conventional methods (Alloing *et al.*, 1998) and transformants were selected using erythromycin (1µg/ml), analyzed by PCR to confirm the presence of the plasmid, and further investigated for the presence of the capsule.

Transmission Electron Microscopy

For ultrastructural analysis of the polysaccharide capsule, bacterial strains
were grown on Columbia agar plates and were prepared following the immune-stabilization method (Jacques *et al.*, 1997). Briefly, bacteria were resuspended in 0.5 ml of PBS pH 7.4 and incubated with 0.1 ml of type 7 antiserum (Statens Serum Institute) for 1 hr at 4°C. After immune-stabilization, bacterial cells were centrifuged at 2500 g for 10 minutes at 4°C and fixed in fixation solution containing 2.5 % (v/v) 7 glutaraldehyde, 0.075 M lysine acetate in cacodylate buffer (0.1 M, pH 7.0) with 0.075% (w/v) ruthenium red (CR buffer) for 2 hrs at 4°C. Controls processed using an unrelated antiserum (type 21), or without antiserum stabilization, were performed simultaneously. All bacterial samples were then washed three times with CR buffer and post-fixed with 1% (w/v) osmium tetroxide in CR buffer for 1 h at RT. Subsequently, samples were washed three times with cacodylate buffer, dehydrated through a graded series of ethanol and embedded in Agar 100 epoxy resin. Ultrathin sections of bacterial samples were placed on formvar-coated copper grids, post-stained with uranyl acetate and lead citrate, and examined with a Philips 208s electron microscope at 80 kV.

Type 7F capsular polysaccharide detection by high-resolution magic angle spinning NMR spectroscopy (HR-MAS NMR)

Bacteria recovered from THYE were inactivated by 1% v/v formaldehyde
treatment and then washed three times with PBS in deuterium oxide (D2O, Sigma-Aldrich, St. Louis, MO, USA).
Approximately 50 µl of compact pellet were inserted in a Kel-F disposable insert for 50 µl volume and then in a 4 mm MAS ZrO2 rotor (Bruker, Madison, WI, USA). Proton HR-MAS NMR experiments were recorded by a Bruker Avance III 400 MHz spectrometer using a Bruker 4-mm HR-MAS probe. The spectra were recorded at 4500 Hz spin rate and 25°C. The 1H spectra were acquired with the combination of a diffusion filter pulse sequence with gradient pulses (diffusion filter 85%), to remove the low molecular weight species free in solution, and a Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence [90-(τ-180-τ)n-acquisition] as T2 filter (76.8 ms), to remove
the broad signals of larger molecular species. 5 mg of purified 7F polysaccharide (Merck & Co., Whitehouse Station, NJ, USA) were solubilized in 0.75 ml of D2O (D2O, Sigma-Aldrich, St. Louis, MO, USA), the solution inserted in 5-mm NMR tube (Wilmad, Vineland, NJ, USA) and proton NMR experiment recorded at 25°C by the same spectrometer, using a 5-mm broadband probe (Bruker, Madison, WI, USA). 1H NMR spectrum was recorded using a standard one-pulse 8 experiment. Both the solid- and liquid-state NMR spectra were collected with 32 k data points over a 10 ppm spectral width. The transmitter was set at the HDO frequency, which was also used as reference signal (4.79 ppm). The TOPSPIN 2.1 software package (Bruker, Madison, WI, USA) was used for data acquisition and processing of all spectra.

Results

S. pneumoniae NT invasive strains AP422 and AP426 contain a 7F capsular locus with point mutations in wchA.

The two invasive *S. pneumoniae* strains AP422 and AP426 were defined NT since serotype testing (confirmed by the Statens Serum Institute, Copenhagen) did not reveal positivity for any capsular type. In order to verify if these strains belonged to a clonal lineage specifically referable to NT pneumococci, they were subjected to MLST analysis. Both AP422 and AP426 were ST191, which, according to a search performed on the *S. pneumoniae* MLST database (www.mlst.net, accessed on 08.18.2011), is mainly (96%) associated with serotype 7F. Using PCR amplification, 7F *cps* was detected in both AP422 and AP426 and then sequenced. Comparison with two reference 7F *cps* sequences (GenBank accession numbers CR931643 and ABFT01000009) revealed the presence of the entire 7F locus with an overall nucleotide identity of 99.8% in both strains. Sequence analysis highlighted the existence of a distinct single point mutation in the *wchA* gene in both strains. With respect to the reference *wchA* nucleotide sequence, AP422 *wchA* showed a deletion of an adenine in position 910 and AP426 *wchA* a G1081T transvertion. Both mutations led to the formation of a premature stop codon associated with the translation of a protein truncated at the C-terminus (Fig. 1). Since the NCBI Conserved Domain algorithm (www.ncbi.nlm.nih.gov/structure/cdd/cdd.shtml) predicts the WchA catalytic domain to be at the C-terminus of the protein (Fig. 1), the truncated WchA proteins coded by AP422 and AP426 are predicted to be enzymatically inactive. In the polysaccharide biosynthesis pathway, WchA (also known as CpsE) is the initial glucose phosphate-transferase, responsible for the linkage of an activated glucose phosphate to the lipid carrier (Pelosi *et al.*, 2005). Since WchA has been demonstrated to be required for the biosynthesis of other capsular types (Pelosi *et al.*, 2005), the mutations identified in *wchA* in both AP422 and AP426 are likely to be responsible for their non-typeability.

```
554/62  1    MDEKGLKIFLAVLQSIIVILLVYFLSFVREAELERSSMVILYLLHFFVFYFSSYGNNFFKRGYLVEFNSTIRYIFFFAIAISVLNFFIAERFSISRRGMV
AP422   1    MDEKGLKIFLAVLQSIIVILLVYFLSFVREAELERSSMVILYLLHFFVFYFSSYGNNFFKRGYLVEFNSTIRYIFFFAIAISVLNFFIAERFSISRRGMV
AP426   1    MDEKGLKIFLAVLQSIIVILLVYFLSFVREAELERSSMVILYLLHFFVFYFSSYGNNFFKRGYLVEFNSTIRYIFFFAIAISVLNFFIAERFSISRRGMV

554/62  101  YFLTLEGISLYLLNFLVKKYWKHVFFNLKNSKKILLLTVTKNMEKVLDKLLESDELSWKLVAVSVLDKSDFQHDKIPVIEKEKIIEFATHEVVDEVFVNL
AP422   101  YFLTLEGISLYLLNFLVKKYWKHVFFNLKNSKKILLLTVTKNMEKVLDKLLESDELSWKLVAVSVLDKSDFQHDKIPVIEKEKIIEFATHEVVDEVFVNL
AP426   101  YFLTLEGISLYLLNFLVKKYWKHVFFNLKNSKKILLLTVTKNMEKVLDKLLESDELSWKLVAVSVLDKSDFQHDKIPVIEKEKIIEFATHEVVDEVFVNL

554/62  201  PGESYDIGEIISRFETMGIDVTVNLKAFDKNLGRNKQIYEMVGLNVVTFSTNFYKTSHVISKRILDICGATIGLILFAIASLVLVPLIRKDGGPAIFAQT
AP422   201  PGESYDIGEIISRFETMGIDVTVNLKAFDKNLGRNKQIYEMVGLNVVTFSTNFYKTSHVISKRILDICGATIGLILFAIASLVLVPLIRKDGGPAIFAQT
AP426   201  PGESYDIGEIISRFETMGIDVTVNLKAFDKNLGRNKQIYEMVGLNVVTFSTNFYKTSHVISKRILDICGATIGLILFAIASLVLVPLIRKDGGPAIFAQT

554/62  301  RIGKNGRHFTFYKFRSMRSDAEAIKEQLMDQNTMQGGMFKIDNDPRVTKIGRFIRKTSLDELPQFWNVFIGDMSLVGTRPPTVDEYVQYTPEQKRRLSFK
AP422   301  RIGKM VDILPFINSVRCGAMLKLSKNS
AP426   301  RIGKNGRHFTFYKFRSMRSDAEAIKEQLMDQNTMQGGMFKIDNDPRVTKIGRFIRKTSLD

554/62  401  PGITGLWQVSGRSKITDFDDVVKLDVAYIDNWTIWKDIEILLKTVKVVFMREGAK
AP422
AP426
```

Fig. 1: Alignment of the WchA amino acids sequences of the reference strain 554/62 (GenBank accession number CR931643) and the two NTPn AP422 and AP426. In both AP422 and AP426 *wchA* is translated as a C-terminal truncated protein. Bold types indicate the catalytic domain as predicted by NCBI Conserved Domain; italic types indicate an alternative C-terminus due to a frame shift in the nucleotide sequence of AP422.

Over-expression of a functional WchA restores the ability of S. pneumoniae to synthesize the 7F capsule.

With the aim to demonstrate that the mutations present in *wchA* were sufficient to account for the lack of capsule in AP422 and AP426, a functional WchA was over-expressed in both NTPn strains. AP422 and AP426 were transformed with the recombinant plasmid pMU1328_*Pc_wchA* to generate the respective recombinant strains AP422∇pMU1328_*Pc_wchA* and AP426∇pMU1328_*Pc_wchA*. The two transformed NTPn yielded a positive reaction with the 7F antiserum in the Quellung reaction. To confirm that WchA over-expression had restored the ability to synthesize a capsule at levels comparable to a wild type capsulated strain, the NTPn isolates along with their WchA over-expressing derivatives were analyzed by TEM and by HR-MAS NMR using the serotype 7F strain AP425 as a positive control. For TEM analysis, the two NTPn strains along with AP425 and the WchA over-expressing derivatives were stabilized with type 7 antiserum before processing for TEM observation. This method was chosen as an alternative of the standard fixation protocol with lysine and ruthenium red (Hammerschmidt *et al.*, 2005); indeed, the latter appeared inadequate to visualize the 7F capsule polysaccharide (data not shown) since it does not contain anionic charged moieties (Moreau *et al.*, 1988). After immune-stabilization a well-preserved dense and thick capsular material was surrounding the AP425 bacterial cells (Fig. 2a), while non-typeable AP422 (Fig. 2b)

and AP426 (Fig. 2c) did not exhibit any characteristic polysaccharide layer on the surface. In WchA over-expressing AP422 and AP426 (Fig. 2d and 2e) TEM examination revealed cells surrounded by a capsular polysaccharide structure comparable to that of AP425, although the capsule of WchA over-expressing AP422 appeared slightly irregular. In order to confirm the data obtained by TEM, proton HR-MAS NMR spectra directly recorded on inactivated cells in the heterogeneous phase was performed. ^{1}H NMR spectrum collected on purified polysaccharide in liquid state clearly revealed specific peaks corresponding to the 7F capsular polysaccharide, such as the anomeric protons at 5.0÷5.6 ppm and the C6 methyl group of rhamnose at 1.2÷1.4 ppm (Fig. 3a). Spectra analysis showed the absence of the 7F specific peaks in the two NTPn and in AP422 and AP426 transformed with an empty pMU1328 plasmid (Fig. 3b) while the above mentioned 7F specific peaks were present both in the capsulated AP425 and in the WchA over-expressing AP422 and AP426 (Fig. 3b). Although signals of other ^{1}H-NMR sensitive molecules expressed on the bacterial surface are partially overlapped with the capsular polysaccharide peaks, the anomeric proton at 5.6 ppm and the C6 methyl group of Rhamnose at 1.2-1.4 ppm fall in spectral windows without other peaks and the assignment results certain.

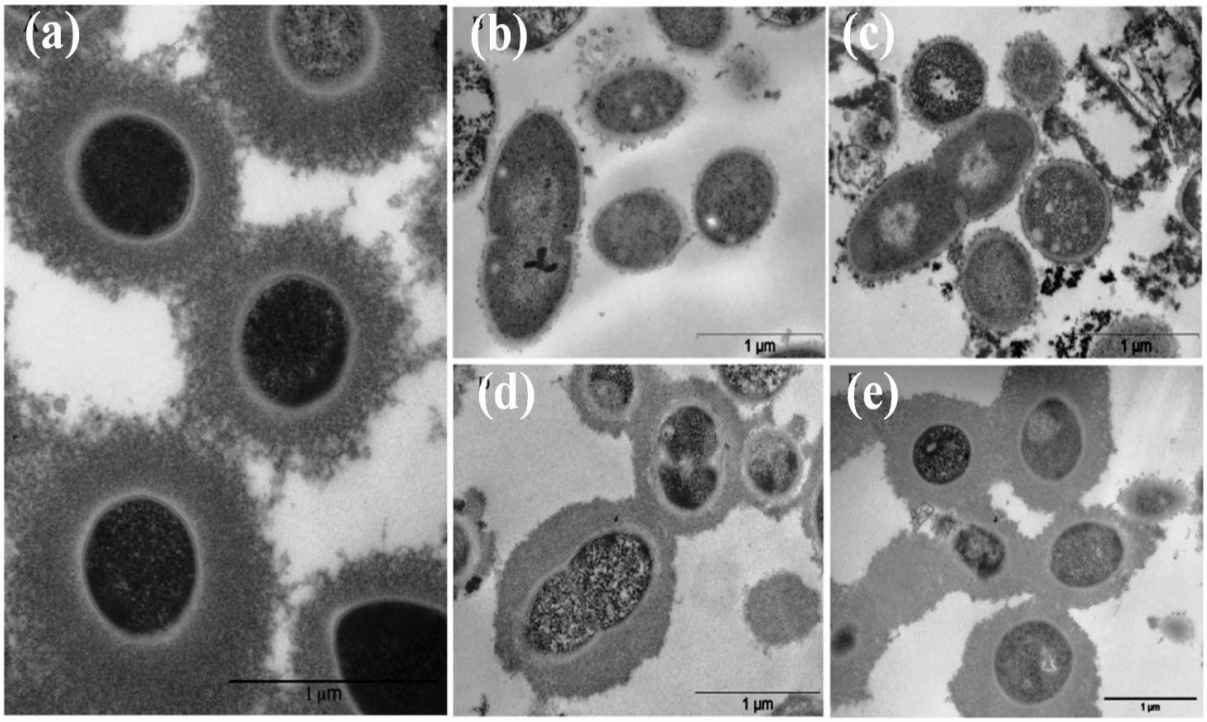
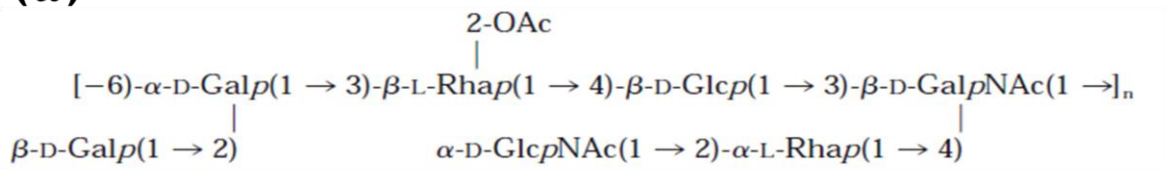


Fig. 2: Transmission Electron Microscopy (TEM) of *S. pneumoniae* strains incubated with type 7 antiserum and stained with ruthenium red. (a): AP425 (serotype 7F); (b): AP422 (NT); (c): AP426 (NT); (d): WchA over-expressing AP422 strain; (e): WchA over-expressing AP426 strain. Over-expression of WchA restored capsule production in both AP422 and AP426.

(a)



(b)

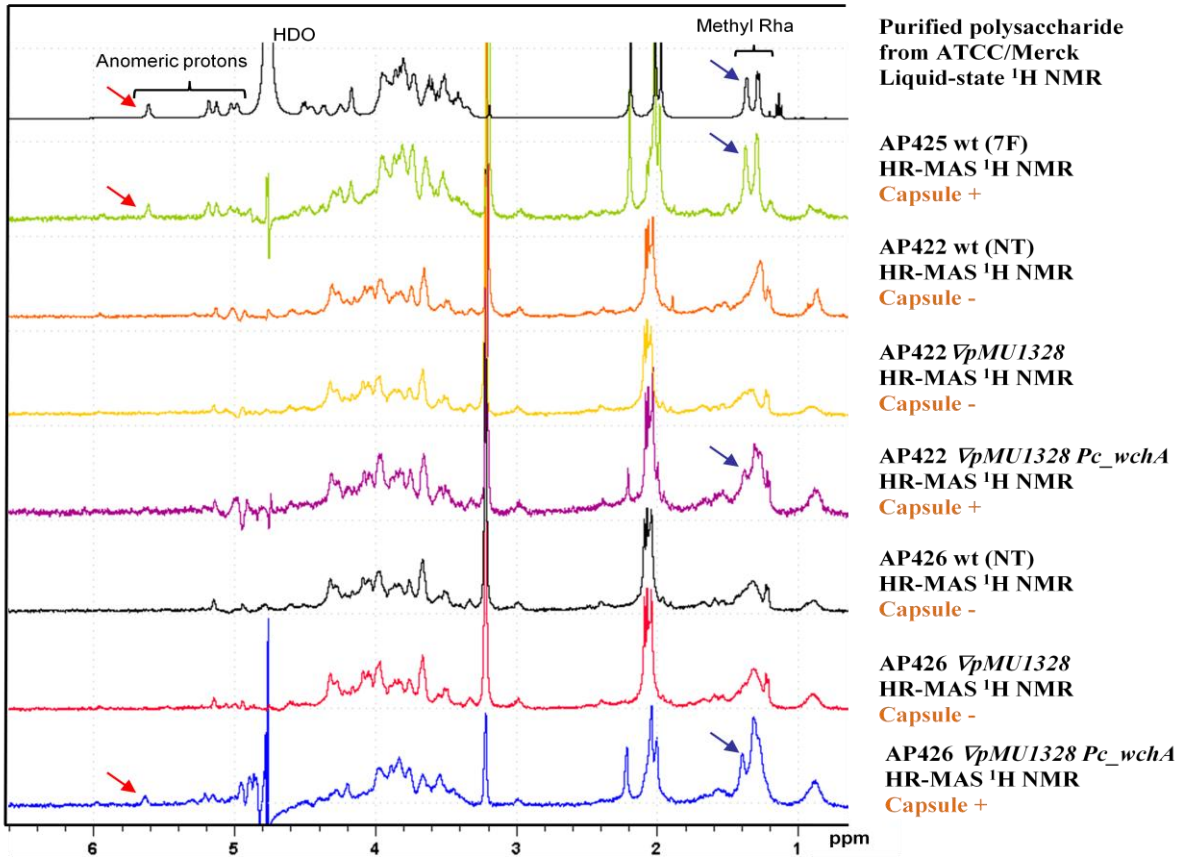


Fig. 3: Proton NMR spectra of 7F purified polysaccharide and *S. pneumoniae* strains. (a) Schematic representation of the chemical structure of 7F polysaccharide. (b) Proton HR-MAS NMR spectra recorded in heterogeneous phase on inactivated bacterial cells. The presence or the absence of the capsule on the bacterial surface, based on spectra analysis, is indicated for each strain. Red arrows indicate anomeric proton at 5.6 ppm, while blue arrows indicate the C6 methyl group of rhamnose at 1.2÷1.4 ppm. 7F purified polysaccharide solubilised in denaturated water is used as a positive control.

Discussion

It is commonly assumed that NTPn lack the polysaccharide capsule, the principal pneumococcal virulence factor that acts by protecting the bacteria during the early phase of infection. Capsule deficiency impairs the capability of NTPn to evade the host's immune defence so that only rarely these strains are able to persist in the blood and cause invasive disease. Indeed, very few cases of invasive diseases due to NTPn have been reported in large epidemiological studies (Beall *et al.*, 2006).

In this study, we have investigated the molecular basis responsible for non-typeability of two NT strains isolated from cases of bacteremic pneumonia in the same hospital during the same year. Molecular analysis of the two invasive NTPn, revealed that these strains were genetically related, since both belonged to ST191 and carried the 7F *cps*. However, they presented distinct point mutations in the *wchA* gene that were responsible for their non-encapsulated phenotype. *wchA* is responsible for the first step of capsule biosynthesis, since it codes for a glycosyltransferase that catalyzes the transfer of the initial sugar to a lipid acceptor, thus initiating the synthesis of the capsule repeat units (Pelosi *et al.*, 2005). Both mutations, introducing a premature stop codon in *wchA*, are expected to produce a glycosyltransferase truncated in its catalytic domain (Pelosi *et al.*, 2005). Re-establishment of a functional *wchA* in both NTPn restored the 7F capsule production as clearly demonstrated by serotyping, TEM and HR-MAS NMR. HR-MAS spectroscopy appeared as a powerful methodology for detection and identification of the bacterial capsule and to confirm its identity. In fact, although precise quantification of the polysaccharide capsule content could not be provided due to the variable number of cells used as input in the analysis, this method clearly gave a semi-quantitative estimation of the specific polysaccharide expressed on the bacterial surface.

Our report is the first describing naturally occurring *wchA* mutations responsible for the NT phenotype in pneumococci. The essentiality of *wchA* for capsular assembly has previously been shown in laboratory mutants of other *S. pneumoniae* serotypes. Insertion-duplication mutagenesis or introduction of a premature stop codon in *wchA* made pneumococci of serotype 19F, 14, 9N, 13 and 15B unable to produce the capsule (Guidolin *et al.*, 1994; Kolkman *et al.*, 1996; Kolkman *et al.*, 1998). Furthermore, sequence duplications within *wchA* was found to be responsible for capsular phase variation in sorbarods-generated pneumococci of serotypes 3, 8 and 37 (Waite *et al.*, 2001; Waite *et al.*, 2003). Hathaway *et al*. reported that in approximately 30% of naturally occurring NTPn the *cps* was present. Non-typeability was assumed to be due to mutations in the *cps* genes, but sequencing was not performed and the gene(s) involved remained unknown (Hathaway *et al.*, 2004).

In conclusion, this study identified two new point mutations, involving *wchA,* responsible for the unencapsulated phenotype in two NTPn isolated from invasive disease. In addition, it provides evidence regarding the key role of *wchA* in serotype 7F capsule biosynthesis process. It remains to be ascertained whether *wchA* is prone to point mutations, leading to the loss of capsule, also in other NTPn strains.

References

Achen, M. G., B. E. Davidson, and A. J. Hillier. 1986. Construction of plasmid vectors for the detection of streptococcal promoters. Gene 45:45-49.

Alloing, G., B. Martin, C. Granadel, and J. P. Claverys. 1998. Development of competence in *Streptococcus pneumoniae*: pheromone autoinduction and control of quorum sensing by the oligopeptide permease. Mol Microbiol 29:75-83.

Andrade, A. L., C. M. Franco, J. Lamaro-Cardoso, M. C. Andre, L. L. Oliveira, A. Kipnis, C. G. Rocha, J. G. Andrade, S. L. Alves, I. H. Park, M. H. Nahm, S. G. Almeida, and M. C. Brandileone. 2010. Non-typeable *Streptococcus pneumoniae* carriage isolates genetically similar to invasive and carriage isolates expressing capsular type 14 in Brazilian infants. J Infect 61:314-322.

Arrecubieta, C., R. Lopez, and E. Garcia. 1994. Molecular characterization of cap3A, a gene from the operon required for the synthesis of the capsule of *Streptococcus pneumoniae* type 3: sequencing of mutations responsible for the unencapsulated phenotype and localization of the capsular cluster on the pneumococcal chromosome. J Bacteriol 176:6375-6383.

Austrian, R. 1976. The quellung reaction, a neglected microbiologic technique. Mt Sinai J Med 43:699-709.

Baldry, S. J., J. Hinds, K. Gould, L. Lambertsen, W. Hanage, M. Antonio, P. Turner, P. Hermans, and S. Bentley. 2009. Presented at the 9th Eur Meet Mol Biol Pneum (Europneumo), Bern, Switzerland.

Beall, B., M. C. McEllistrem, R. E. Gertz, Jr., S. Wedel, D. J. Boxrud, A. L. Gonzalez, M. J. Medina, R. Pai, T. A. Thompson, L. H. Harrison, L. McGee, and C. G. Whitney. 2006. Pre- and postvaccination clonal compositions of invasive pneumococcal serotypes for isolates collected in the United States in 1999, 2001, and 2002. J Clin Microbiol 44:999-1017.

Bentley, S. D., D. M. Aanensen, A. Mavroidi, D. Saunders, E. Rabbinowitsch, M. Collins, K. Danohoe, D. Harris, L. Murphy, P. R. Reeves, J. Parkhill, and B. G. Spratt. 2006. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes., p. E31. PLoS Genet., vol. 2.

Brown, E. J., S. W. Hosea, and M. M. Frank. 1983. The role of antibody and complement in the reticuloendothelial clearance of pneumococci from the bloodstream. Rev Infect Dis 5 Suppl 4:S797-805.

Calix, J. J., and M. Nahm. 2010. A new pneumococcal serotype, 11E, has a variably inactivated *wcjE* gene. J Infec Dis 202.

Carvalho, M. G., A. G. Steigerwalt, T. Thompson, D. Jackson, and R. R. Facklam. 2003. Confirmation of nontypeable *Streptococcus pneumoniae*-like organisms

isolated from outbreaks of epidemic conjunctivitis as *Streptococcus pneumoniae*. J Clin Microbiol 41:4415-4417.

Gentile, M. A., S. Melchiorre, C. Emolo, M. Moschioni, C. Gianfaldoni, L. Pancotto, I. Ferlenghi, M. Scarselli, W. Pansegrau, D. Veggi, M. Merola, F. Cantini, P. Ruggiero, L. Banci, and V. Masignani. 2011. Structural and functional characterization of the *Streptococcus pneumoniae* RrgB pilus backbone D1 domain. J Biol Chem 286:14588-14597.

Gherardi, G., F. D'Ambrosio, M. Monaco, R. Camilli, L. De Florio, F. D'Ancona, G. Dicuonzo, and A. Pantosti. 2009. Population structure of invasive *Streptococcus pneumoniae* isolates in Italy prior to the implementation of the 7-valent conjugate vaccine (1999-2003). Eur J Clin Microbiol Infect Dis 28:99-103.

Gladstone, R. A., J. M. Jefferies, S. N. Faust, and S. C. Clarke. 2011. Continued control of pneumococcal disease in the UK - the impact of vaccination. J Med Microbiol 60:1-8.

Guidolin, A., J. K. Morona, R. Morona, D. Hansman, and J. C. Paton. 1994. Nucleotide sequence analysis of genes essential for capsular polysaccharide biosynthesis in *Streptococcus pneumoniae* type 19F. Infect Immun 62:5384-5396.

Hammerschmidt, S., S. Wolff, A. Hocke, S. Rosseau, E. Muller, and M. Rohde. 2005. Illustration of pneumococcal polysaccharide capsule during adherence and invasion of epithelial cells. Infect Immun 73:4653-4667.

Hanage, W. P., T. Kaijalainen, A. Saukkoriipi, J. L. Rickcord, and B. G. Spratt. 2006. A successful, diverse disease-associated lineage of nontypeable pneumococci that has lost the capsular biosynthesis locus. J Clin Microbiol 44:743-749.

Hathaway, L. J., P. Stutzmann Meier, P. Battig, S. Aebi, and K. Muhlemann. 2004. A homologue of *aliB* is found in the capsule region of nonencapsulated *Streptococcus pneumoniae*. J Bacteriol 186:3721-3729.

Jacques, M., and M. Gottschalk. 1997. Use of monoclonal antibodies to visualize capsulare material of bacterial pathogens by conventional electron microscopy. Microsc Microanal 3:234-238.

Kolkman, M. A., D. A. Morrison, B. A. Van Der Zeijst, and P. J. Nuijten. 1996. The capsule polysaccharide synthesis locus of streptococcus pneumoniae serotype 14: Identification of the glycosyl transferase gene cps14E. J Bacteriol 178:3736-3741.

Kolkman, M. A., B. A. van der Zeijst, and P. J. Nuijten. 1998. Diversity of capsular polysaccharide synthesis gene clusters in *Streptococcus pneumoniae*. J Biochem 123:937-945.

Llull, D., R. Munoz, R. Lopez, and E. Garcia. 1999. A single gene (*tts*) located outside the cap locus directs the formation of *Streptococcus pneumoniae* type 37 capsular polysaccharide. Type 37 pneumococci are natural, genetically binary strains. J Exp Med 190:241-251.

Moreau, M., J. C. Richards, M. B. Perry, and P. J. Kniskern. 1988. Application of high-resolution n.m.r. spectroscopy to the elucidation of the structure of the specific capsular polysaccharide of *Streptococcus pneumoniae* type 7F. Carbohydr Res 182:79-99.

Moschioni, M., C. Donati, A. Muzzi, V. Masignani, S. Censini, W. P. Hanage, C. J. Bishop, J. N. Reis, S. Normark, B. Henriques-Normark, A. Covacci, R. Rappuoli, and M. A. Barocchi. 2008. Streptococcus pneumoniae contains 3 rlrA pilus variants that are clonally related. J Infect Dis 197:888-896.

Park, I. H., D. G. Pritchard, R. Cartee, A. Brandao, M. C. Brandileone, and M. H. Nahm. 2007. Discovery of a new capsular serotype (6C) within serogroup 6 of *Streptococcus pneumoniae*. J. Clin. Microbiol. 45:1225-1233.

Pelosi, L., M. Boumedienne, N. Saksouk, J. Geiselmann, and R. A. Geremia. 2005. The glucosyl-1-phosphate transferase WchA (Cap8E) primes the capsular polysaccharide repeat unit biosynthesis of *Streptococcus pneumoniae* serotype 8. Biochem Biophys Res Commun 327:857-865.

Sa-Leao, R., A. S. Simoes, S. Nunes, N. G. Sousa, N. Frazao, and H. de Lencastre. 2006. Identification, prevalence and population structure of non-typable *Streptococcus pneumoniae* in carriage samples isolated from preschoolers attending day-care centres. Microbiology 152:367-376.

Waite, R. D., D. W. Penfold, J. K. Struthers, and C. G. Dowson. 2003. Spontaneous sequence duplications within capsule genes cap8E and tts control phase variation in *Streptococcus pneumoniae* serotypes 8 and 37. Microbiology 149:497-504.

Waite, R. D., J. K. Struthers, and C. G. Dowson. 2001. Spontaneous sequence duplication within an open reading frame of the pneumococcal type 3 capsule locus causes high-frequency phase variation. Mol Microbiol 42:1223-1232.

Xu, Q., R. Kaur, J. R. Casey, V. Sabharwal, S. Pelton, and M. E. Pichichero. 2011. Nontypeable *Streptococcus pneumoniae* as an otopathogen. Diagn Microbiol Infect Dis 69:200-204.

Acknowledgements

This is the end…the end of another part of my life!

Direi che questa parte vale la pena scriverla in italiano…

Siamo alla fine…chi l'avrebbe mai detto, eppure sono qui che scrivo questa pagina, i Ringraziamenti. Sono davvero arrivata al momento di dover mettere "il punto" e girare pagina, iniziare un nuovo capitolo.

É stata un'esperienza unica, ricca di ogni cosa, di nozioni acquisite, tecniche imparate, esperienze vissute, power point preparati, documenti inviati ma anche e soprattutto fatta di incontri, di crescita personale, di confronti, di decisioni importanti, di tanti ostacoli affrontati e superati con l'aiuto di coloro che mi sono sempre stati accanto, persone speciali che porterò sempre con me e che spero non deluderò mai.

Monica che dire…il nostro primo incontro fu particolare, schietto e sincero. Le prime parole che dicesti furono: "Dovrai essere preparata al fatto che qui ci sono parecchie cose da fare, diverse tra loro. Ci vuole organizzazione e capacità di lavorare su più cose". Il primo impatto non fu dei migliori ma poi conoscendoti ho apprezzato la persona meravigliosa che sei. Una preparazione scientifica immensa, senza eguali. Sei riuscita a trasmettermi e ad insegnarmi tutto ciò che sai, non avrei potuto desiderare di meglio. Poi, come succede nelle migliori famiglie, abbiamo avuto degli scontri, incomprensioni ma solo perché volevamo il meglio l'una per l'altra, come "due sorelline".
Insomma una combinazione perfetta, fatta di lavoro e amicizia, complicità e indipendenza, divertimento e serietà, rispetto per i miei momenti di silenzio che di certo non sono mancati. Insomma…cosa vuoi di più?

"La mia Gabri"…per lei non ci sono parole, per il rapporto, il nostro rapporto, non ci sono parole. Il legame che ci unisce è qualcosa davvero di raro, unico, assolutamente non scontato e facile da poter trovare.
Ci siamo incontrate dopo la prima settimana di università e non ci siamo più lasciate. Nove anni di vita vissuti insieme, vicine, gomito a gomito, senza mai un litigio, direi quasi assurdo, da non credere! Eppure è così, ci siamo sempre confrontate, aiutate ma mai accusate, siamo state sempre il sostegno l'una dell'altra, sincere più che mai, leali, sempre presenti.
Ora ci siamo separate e non sai, non puoi immaginare, cosa è significato per me. Sento un vuoto, un'assenza forte, a volte enorme…mi sento sola.
Anche se so perfettamente che ci saremo sempre l'una per l'altra, anche a mille miglia di distanza; perché lo sai Ga, le cose VERE, profonde non scompariranno mai.

Michèle…il "mio capo"…magari fossero tutti così i capi!
Sempre solare, sorridente, disponibile, con la quale poter parlare di tutto, insomma un'amica.
Grazie per avermi dato la possibilità di vivere quest'esperienza decidendo di prendermi nel tuo gruppo.
Grazie per aver supportato i miei studi e la mia ricerca, che con Monica avete studiato a tavolino per permettermi di raggiungere obiettivi elevati che mi garantissero un ottimo livello di ricerca.

Grazie per l'esperienza americana che hai permesso che io vivessi. Sicuramente è stata dura ma mi ha dato così tanto, mi ha insegnato tanto, soprattutto nel saper interagire con gli altri.
Grazie per l'aiuto che continui a darmi.
Grazie per aver creduto in me.

Infine ma non ultima, la Mia Famiglia…sempre presente!
La mia mamma e il mio papà, due fari per me. La mia "sorellina", una certezza. Strano accorgersi come crescendo senti un richiamo a ciò che sono le tue radici, la tua famiglia…eh sì la vita è così.
Anche in questo caso penso di essere molto fortunata, aver avuto la possibilità di poter crescere e vivere la mia vita in una famiglia eccezionale.
Vorrei ringraziare loro per la tranquillità in cui mi hanno sempre permesso di vivere, per l'educazione che mi hanno dato e per avermi fatto capire cosa è davvero importante nella vita e cosa no.
Vorrei ringraziarli per tutto il sostegno psicologico che mi hanno dato nei sei mesi in America, se non ci fossero stati loro non so come avrei fatto!
Vorrei ringraziarli per ciò che continuano a fare per me, un lavoro costante di sostegno.
Grazie grazie grazie!

Ah dimenticavo l'ultima arrivata…Morena.
E' stata una scoperta meravigliosa, ha portato allegria, spontaneità, simpatia…la cosiddetta napoletaneità!
Grazie per ciò che hai fatto per me, per tutte le dinamiche che mi hai permesso di capire grazie alle nostre mille chiacchierate.
Grazie per essermi stata vicino in un periodo davvero buio per me.
Grazie per ciò che sei, non cambiare mai.

E poi…Grazie a Colui che c'è sempre, che a distanza o al mio fianco è capace di farmi sentire la sua presenza, di farmi sentire protetta.
Un punto di riferimento, qualcosa di insostituibile nella mia vita.

Ehi ma attenzione…non pensate mai di potervi liberare di me!

Sara

ORIGINAL ARTICLE 10.1111/j.1469-0691.2011.03711.x

Assessment of *Streptococcus pneumoniae* pilus islet-1 prevalence in carried and transmitted isolates from mother–infant pairs on the Thailand–Burma border

P. Turner[1,2,3], S. Melchiorre[4], M. Moschioni[4], M. A. Barocchi[4], C. Turner[1,2,3], W. Watthanaworawit[1,2], N. Kaewcharernnet[1,2], F. Nosten[1,2,3] and D. Goldblatt[5]

1) *Shoklo Malaria Research Unit, Mae Sot,* 2) *Mahidol-Oxford Tropical Medicine Research Unit, Bangkok, Thailand,* 3) *Centre for Tropical Medicine, University of Oxford, Oxford, UK,* 4) *Novartis Vaccines and Diagnostics s.r.l., Siena, Italy and* 5) *Immunobiology Unit, Institute of Child Health, University of London, London, UK*

Abstract

Streptococcus pneumoniae pilus islet-1 (PI-1)-encoded pilus enhances *in vitro* adhesion to the respiratory epithelium and may contribute to pneumococcal nasopharyngeal colonization and transmission. The pilus subunits are regarded as potential protein vaccine candidates. In this study, we sought to determine PI-1 prevalence in carried pneumococcal isolates and explore its relationship with transmissibility or carriage duration. We studied 896 pneumococcal isolates collected during a longitudinal carriage study that included monthly nasopharyngeal swabbing of 234 infants and their mothers between the ages of 1 and 24 months. These were cultured according to the WHO pneumococcal carriage detection protocol. PI-1 PCR and genotyping by multilocus sequence typing were performed on isolates chosen according to specific carriage and transmission definitions. Overall, 35.2% of the isolates were PI-1-positive, but PI-1 presence was restricted to ten of the 34 serotypes studied and was most frequently associated with serotypes 19F and 23F; 47.5% of transmitted and 43.3% of non-transmitted isolates were PI-1-positive (OR 1.2; 95% CI 0.8–1.7; p 0.4). The duration of first-ever infant pneumococcal carriage was significantly longer with PI-1-positive organisms, but this difference was not significant at the individual serotype level. In conclusion, PI-1 is commonly found in pneumococcal carriage isolates, but does not appear to be associated with pneumococcal transmissibility or carriage duration.

Keywords: Carriage, carriage duration, colonization, PI-1, pilus-1, *Streptococcus pneumoniae*, transmission
Original Submission: 8 July 2011; **Revised Submission:** 20 October 2011; **Accepted:** 25 October 2011
Editor: J.-L. Mainardi

Clin Microbiol Infect

Corresponding author: P. Turner; Shoklo Malaria Research Unit, PO Box 46, 68/30 Ban Toong Road, Mae Sot 63110, Thailand.
E-mail: pault@tropmedres.ac

Introduction

Streptococcus pneumoniae is a significant global pathogen [1]. Additionally, pneumococcus is a nasopharyngeal commensal, and infants are frequently colonized [2,3]. Various cell surface components, including pneumococcal surface adhesin A, choline-binding protein A, pneumococcal serine-rich repeat protein, and pneumococcal adherence-virulence factor A, have been shown to contribute to nasopharyngeal adherence and colonization, but their mechanisms of action remain incompletely understood [3–7]. The RrgA subunit of the surface exposed pilus-1 filamentous structure also enhances pneumococcal adherence to respiratory epithelial cells *in vitro* [8,9]. Pneumococcal pilus-1 is encoded by the pilus islet-1 (PI-1; *rlrA* islet), and is composed of three subunits: RrgA, RrgB, and RrgC [9–11]. As immunization with pilus antigens is protective against lethal intraperitoneal challenge in a mouse model, pilus subunits are regarded as potential candidates for inclusion in a protein-based pneumococcal vaccine [12]. However, PI-1 presence is not universal: studies of predominantly invasive pneumococci have found that PI-1 is present in isolates from a limited number of serotypes

(particularly those included in the pneumococcal seven-valent conjugate vaccine (PCV7)) and that its presence correlates with genotype by multilocus sequence typing (MLST) [13–18]. Given the prevalence of asymptomatic colonization, and the fact that the serotype/genotype structure of invasive pneumococci is not stable [19], further work to describe the prevalence and function of pilus-1 in pneumococcal carriage strains is warranted.

In this study, our aims were to: (i) determine PI-1 prevalence in pneumococcal carriage isolates from infants and their mothers; (ii) determine which PI-1-positive clones were associated with carriage in the study region; and (iii) explore the possible functions of the pilus *in vivo*, and in particular evaluate whether PI-1 presence had an effect on pneumococcal transmission or carriage duration.

Materials and Methods

Pneumococcal isolates

The isolates characterized in this work were collected from 2007 to 2010, during a longitudinal carriage and pneumonia study of 965 infants at Maela refugee camp, Thailand. Maela was established in 1984, and *c.* 40 000 Burmese refugees, predominantly of Karen ethnicity, live in a 4-km^2 area. Pneumococcal vaccines are not available in the camp. Nested within the study was a subcohort of 234 mother–infant pairs who were studied in more detail: between 1 and 24 months of age, these infants had a nasopharyngeal swab (NPS) and blood taken at monthly surveillance visits, and their mothers had an NPS taken at the same time. NPSs were processed according to the WHO pneumococcal colonization protocol, as previously described [20,21]. All morphologically distinct pneumococcal isolates from each swab culture were serotyped by latex agglutination [22]. Non-typeable (NT) isolates, either morphologically typical pneumococcal colonies (probably encapsulated organisms) or rough colonies (probably non-encapsulated organisms) that were non-reactive with typing antisera, were confirmed by bile solubility and absent capsular swelling with Omniserum (Statens Serum Institute, Hillerod, Denmark). Serotype 6C was identified by PCR [23].

Mother and infant pneumococcal carriage patterns for the first 12 months of follow-up were reviewed, and isolates were chosen for further analysis according to the definitions described below.

Definition of transmission and carriage

Transmission of a pneumococcus was defined as the presence of an identical pneumococcal serotype and MLST genotype in both mother and infant at the same visit ('concordant transmission') or when the pneumococcus was isolated from the mother and the infant during a 'transmissibility episode' but not at the same visit ('discordant transmission'). A non-transmitted pneumococcus was a carried serotype that never appeared in the other member of the mother–infant pair during the transmissibility episode as defined below (Fig. S1). Transmissibility episodes were identified by carriage within a mother–infant pair when there was the same pneumococcal serotype cultured from two or more NPSs (from mother and/or infant), separated by two or fewer NPSs negative for that serotype. For serotypes 1 and 5, which are known to be carried for very short durations, a single positive NPS could define non-transmission.

For the determination of carriage duration in infants, a pneumococcal acquisition was defined as midpoint between the last negative swab and the first positive swab for a given serotype. Termination of the carriage episode was similarly defined as the midpoint between the first of two consecutive negative swabs and the last positive swab for the serotype.

PI-1 detection

PI-1 presence was determined as previously described [14]. Briefly, PCRs were performed directly from bacteria, with the primers listed in Table S1. Primers were designed on conserved regions on the boundaries of PI-1 (459for and 470rev) and within PI-1 (P01rev, P11for, P08for, and P08rev).

MLST

MLST was performed as previously described [24]. Briefly, PCR amplifications were performed directly from the bacteria with the standard primer pairs. Sequences were obtained on both strands with an ABI 3730xl DNA Analyzer (Life Technologies Corporation, Carlsbad, CA, USA). Sequence type (ST) was determined by use of the MLST website (http://spneumoniae.mlst.net). eBURST (http://spneumoniae.mlst.net/eburst/) was run with default settings on the entire MLST database, and each ST was assigned to a clonal complex (CC) [25]. CCs were named in accordance with the ST number of the founder predicted by eBURST.

Statistical analysis

Statistical analyses were carried out with STATA 10.1 (StataCorp, College Station, TX, USA). The chi-squared test, Fisher's exact test and ORs were used to compare proportions. Carriage duration was estimated by survival analysis methods, with the log-rank test being used to compare groups [2].

Study ethics

Ethical approval for the carriage study was granted by Mahidol University, Thailand and Oxford University, UK.

Results

Pneumococcal isolates and PI-1 prevalence

In the first 12 months of follow-up, 4921 surveillance NPS were collected from the 234 mother–infant pairs (84.7% of expected). In total, 2497 isolates were cultured, and 896 (35.9%) of these were included in the current work. All isolates meeting the criteria described below were included, resulting in 34 serotypes being represented in the isolate selection. These serotypes accounted for 90.1% (2251/2497) of isolates in the entire cohort isolate collection.

We successfully determined the PI-1 status of 887/896 (98.9%) isolates: 35.2% (312/887) were positive. PI-1-positive isolates were restricted to ten serotypes: 4, 6A, 6B, 9, 14, 19F, 19A, 23F, 33C, and NT (Fig. 1).

PI-1 and transmission

To determine whether PI-1 presence had an effect on pneumococcal transmission within the mother–infant pairs, a panel of isolates was selected and analysed by both PI-1 PCR and MLST (to define genotype). We elected to study serotypes contained in, or related to, current conjugate vaccines (PCV13 + 6C), as well as NT isolates, which were overrepresented in our population (the commonest pneumococcal 'type' identified from mother swabs, and the third commonest in infants). In the first year of follow-up, these serotypes accounted for 69.6% and 59.8% of isolates carried by infants and mothers, respectively. In the case of a potentially transmitted pneumococcal serotype, i.e. isolates carried by both mother and infant at the same visit, or at sequential visits, single isolates from both mother and infant were selected for study. For non-transmitted isolates, i.e. isolates carried by only the mother or infant, a single isolate was studied. Additionally, we analysed every isolate from carriage episodes of all serotypes in eight mother–infant pairs to determine the clonality of pneumococcal carriage.

Of the 489 isolates selected for the primary analysis, PI-1 PCR was uninterpretable in six, resulting in 483 analysable isolates. Overall, 219/483 (45.3%) of these were PI-1 positive. PI-1 was found in NT isolates and in 6/14 serotypes studied, all of which, apart from one 19A isolate, were PCV7 serotypes (Table 1). PI-1 was found most frequently in 19F (91%) and 23F (77%) isolates, although these serotypes, according to the MLST and eBURST analyses, were dominated by a single CC each (19F, CC271; 23F, CC802). Overall, PI-1-positive isolates were members of a restricted group of CCs (Fig. 2).

With serotyping of a representative isolate alone, 253/483 (52.4%) isolates would have been classified as transmitted (153 concordant; 100 discordant) and 230/483 (47.6%) as non-transmitted. With a combination of MLST and serotype, nine concordant and eight discordant 'transmitted' isolates (17/483; 3.5%) were reclassified as non-transmitted, resulting

TABLE 1. Pneumococcal serotype distribution and isolate transmission category in mother–infant pairs (pneumococcal 13-valent conjugate vaccine serotypes plus 6C and nontypeable (NT))

Serotype	Total, *N*	Transmitted[a], *N* (% within serotype)	PI-1 present, *N* (% within serotype)
1	12	4 (33.3)	0 (0.00)
3	11	4 (36.4)	0 (0.00)
4	4	2 (50.0)	0 (0.00)
5	10	2 (20.0)	0 (0.00)
6A	16	6 (37.5)	0 (0.00)
6B	54	27 (50.0)	29 (53.7)
6C	15	4 (26.7)	0 (0.00)
7F	5	2 (40.0)	0 (0.00)
9V	8	4 (50.0)	5 (62.5)
14	46	28 (60.9)	15 (32.6)
18C	7	4 (57.1)	0 (0.00)
19F	89	50 (56.2)	81 (91.0)
19A	14	4 (28.6)	1 (7.1)
23F	82	38 (46.3)	63 (76.8)
NT	110	57 (51.8)	25 (22.7)
Total	483	236 (48.9)	219 (45.3)

PI-1, pilus islet-1.
[a]Concordant and discordant transmission combined.

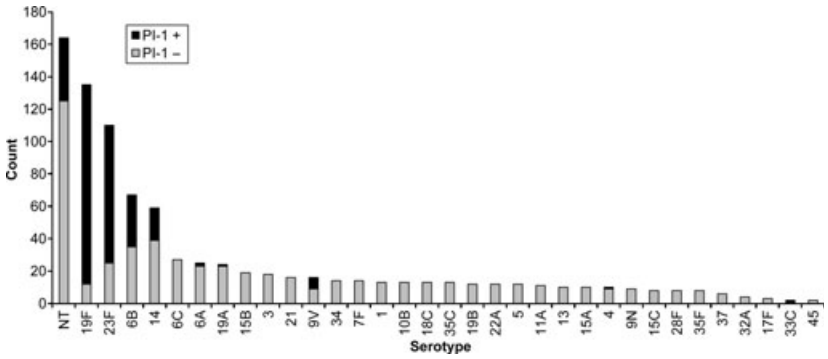


FIG. 1. Pilus islet-1 (PI-1) presence by serotype in 887 carried pneumococcal isolates (includes all commonly carried serotypes).

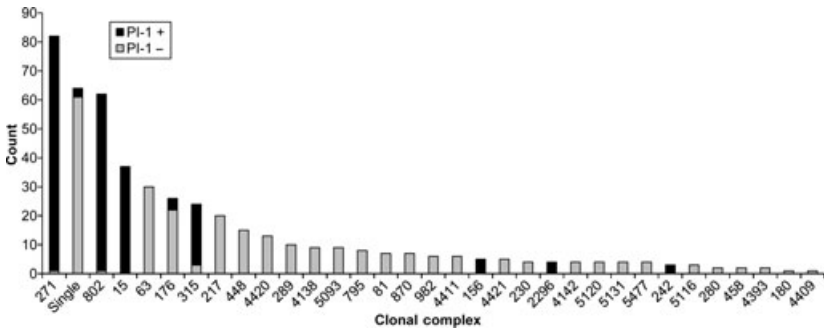


FIG. 2. Pilus islet-1 (PI-1) presence by clonal complex in 483 carried pneumococcal isolates (pneumococcal 13-valent conjugate vaccine serotypes plus 6C and non-typeable (NT) only).

TABLE 2. First pneumococcal carriage episode duration (six commonest serotypes)

Serotype	Number of episodes	Median carriage duration (days) (95% CI)	Mean carriage duration (days) (95% CI)
6B	13	121 (90–153)	119 (93–145)
14	8	62 (30–151)	86 (45–128)
19F	21	213 (63–243)	231 (154–308)
23F	21	184 (62–277)	176 (124–229)
35F	9	121 (30–180)	124 (66–182)
NT	33	31 (31–61)	72 (44–99)

NT, non-typeable.

in final totals of 236/483 (48.9%) transmitted and 247/483 (51.1%) non-transmitted isolates. PI-1 presence was not correlated with transmission: 47.5% (112/236) of transmitted and 43.3% (107/247) of non-transmitted isolates were PI-1 positive, with an OR of 1.2 (95% CI 0.8–1.7; p 0.4), and no significant differences were seen at the individual serotype level.

We analysed an additional 91 isolates, from seven mother–infant pairs and one mother–twin infant unit, to confirm that sequential isolates of the same serotype in an individual, and isolates of the same serotype in a mother–infant pair, were indeed identical. In all mother–infant pairs, each serotype carried was represented by a single ST per carriage/transmission episode, with three exceptions (Table S2).

PI-1 and first pneumococcal carriage episode duration in infants

To analyse the correlation between PI-1 presence and the duration of infant carriage, we determined the PI-1 status of 316 isolates from the first carriage episodes of all common serotypes (defined as serotypes with at least ten carriage episodes). These episodes included first episodes of carriage of each serotype, not only an infant's first ever pneumococcal carriage episode. We inferred the PI-1 status for the entire carriage episode from the PI-1 PCR result of a single isolate.

Given the well-described association of shorter carriage in individuals with prior pneumococcal exposure, we subsequently focused our analysis only on first-ever episodes of carriage (i.e. one per infant) [2]. From the 316 carriage episodes selected, we identified 216 analysable first-ever carriage episodes: PI-1 PCR data were available for at least one isolate in 180 of these (90 pneumococcal 13-valent conjugate vaccine serotypes, 58 non-vaccine serotypes, and 32 NT pneumococci). Median and mean durations of the first pneumococcal carriage episode were 63 days (95% CI 61–91) and 105 days (95% CI 91–119), respectively, but varied considerably by serotype (Table 2). On comparison of the 180 carriage episodes where PI-1 PCR results were available, PI-1-positive carriage episodes (51/180; 28.3%) were significantly longer than those associated with a PI-1-negative organism (median 152 days (95% CI 93–213) vs. 61 days (95% CI 57–90); mean 177 days (95% CI 137–218) vs. 84 days (95% CI 72–97); p <0.0001). However, the analysis of carriage duration was confounded by serotype. In our study, serotypes 19F and 23F had the longest duration of carriage, and were both predominantly PI-1-positive. To analyse the relative contributions of serotype and PI-1 to carriage duration, we fitted a Cox regression model and found that, on stratification of carriage by serotype, PI-1 presence was not associated with a significant change in carriage duration (hazard ratio 0.7 (95% CI 0.4–1.2; p 0.2)).

Discussion

This study is the largest investigation of the prevalence and possible function of the PI-1-encoded pilus in carriage isolates of *S. pneumoniae*. It also provides pneumococcal strain data from Southeast Asia, a densely populated region that is under-represented in the pneumococcal carriage and disease epidemiology literature. As serotypes are commonly used as the basis for defining transmission, we were interested in the

frequency with which a molecular analysis may confound the analysis of transmission. By genotyping, discrepancies were discovered in only 3.5% of transmission events classified by serotype. We also demonstrated that pneumococci of the same serotype are predominantly clonal within a discrete carriage episode. We therefore concluded that serotype can be used to define transmission events and carriage episodes with relatively high confidence, but that the inclusion of genotyping is vital to ensure complete accuracy.

Because of the selection criteria for the isolates included in the study, we cannot describe the overall pilus prevalence for the population. However, in the 887 carriage isolates analysed, the pilus prevalence was 35.2%. Among the 34 serotypes analysed, we found that PI-1 was present only in NT isolates and nine other serotypes: 65% of PCV7 isolates were PI-1-positive, as compared with 9% of non-vaccine serotypes. Interestingly, Basset *et al.*, examining nasopharyngeal and invasive pneumococcal isolates from the American Indian collection, also found that PCV7 strains were significantly more likely to be PI-1-positive than non-vaccine serotypes, but that the overall proportion of PI-1-positive isolates was slightly lower [13]. They, along with other authors, demonstrated that PI-1-positive strains were contained within a small number of CCs [14–16]. Indeed, in the current study, we demonstrated that PI-1-positive isolates were clustered predominantly within four dominant CCs (CC15, CC271, CC315, and CC802). The strong association between serotype/CC and pilus presence, along with strain selection criteria, and the different regional distribution of the clones, may explain the variability in PI-1 prevalence between different studies.

We did not demonstrate a pilus-attributable effect on pneumococcal transmissibility. However, there are limitations to our study that may be important confounders for this analysis. The carriage study was carried out in a densely populated refugee camp where 13% of the population are <5 years old, and there is likely to be frequent transmission of nasopharyngeal organisms both within families and between members of the community. Despite the use of a combination of serotype and ST to increase the accuracy of our transmissibility categories, we cannot exclude the possibility that 'non-transmitted' strains were effectively transmitted between mother or infant and others but not detected by us. Several studies have documented the clustering of pneumococcal serotypes and genotypes within families, which highlights the difficulty of assigning a definitive 'non-transmitted' label to isolates collected from an incomplete household group [26,27]. However, we focused on mother–infant transmission, because we felt that the absence of a strain in one member of this pair would be the best marker for relative non-transmissibility in the early months of life. In addition, we included both concordant and discordant time-point pneumococcal serotype/ST identifications in the mother–infant pair as 'transmitted', as long as they occurred no more than 2 months apart. However, the study sampling frequency may have been too low to demonstrate transmission of serotypes carried for very short durations. Also, in the presence of multiple serotype carriage, a common occurrence in infancy and one that is underestimated by standard culture protocols [21], a particular serotype may become undetectable for a period of time before re-emerging as the dominant serotype, and this may result in incorrect categorization regarding transmissibility.

Although, in the crude analysis, PI-1 presence was associated with longer first pneumococcal carriage episodes in infants, we could find no significant association between carriage duration and PI-1 at the individual serotype level. This is likely to be the result of the low numbers of carriage episodes of individual serotypes (resulting in wide CIs around the carriage duration estimates) and the restricted number of clones within each serotype. We could have increased the number of carriage episodes included in our analysis by looking at all carriage episodes of each serotype rather than restricting our investigation to each infant's first-ever carriage episodes, but this would have introduced other confounding factors, such as the impact of previous carriage and immune factors on subsequent carriage episode duration. Therefore, it is possible that, as a result the study sample size, a small effect of pilus on either transmission or carriage duration may have been missed.

In conclusion, we found that Southeast Asian pneumococcal carriage isolates had a similar pilus prevalence to to that in previously described strain collections, which is helpful in the ongoing assessment of likely global coverage of a pilus subunit-containing vaccine. Despite its known role in pneumococcal attachment, we could not determine a clear impact of PI-1 presence on transmissibility or carriage duration.

Acknowledgements

The authors are grateful to all of the study participants and the SMRU staff at the Maela clinic and Mae Sot laboratories.

Transparency Declaration

This work was supported by Novartis and the Wellcome Trust (grant numbers 083735 (P. Turner) and 077166/Z/05 (C. Turner and F. Nosten)). M. A. Barocchi, M. Moschioni

and S. Melchiorre are employees of Novartis Vaccines and Diagnostics. There are no other potential conflicts to report. This work was presented in part as a poster at the 7th International Symposium on Pneumococci and Pneumococcal Diseases, Tel Aviv, Israel, March 2010 (Abstract 477).

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Pneumococcal serotype transmission examples.

Table S1. PI-1 PCR primer set.

Table S2. Detailed study of pneumococcal carriage in the first year of life in eight mother-infant pairs.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

1. O'Brien KL, Wolfson LJ, Watt JP *et al.* Burden of disease caused by *Streptococcus pneumoniae* in children younger than 5 years: global estimates. *Lancet* 2009; 374: 893–902.
2. Hill PC, Cheung YB, Akisanya A *et al.* Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian infants: a longitudinal study. *Clin Infect Dis* 2008; 46: 807–814.
3. Bogaert D, De Groot R, Hermans PW. *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. *Lancet Infect Dis* 2004; 4: 144–154.
4. Rajam G, Anderton JM, Carlone GM, Sampson JS, Ades EW. Pneumococcal surface adhesin A (PsaA): a review. *Crit Rev Microbiol* 2008; 34: 163–173.
5. Rosenow C, Ryan P, Weiser JN *et al.* Contribution of novel choline-binding proteins to adherence, colonization and immunogenicity of *Streptococcus pneumoniae*. *Mol Microbiol* 1997; 25: 819–829.
6. Shivshankar P, Sanchez C, Rose LF, Orihuela CJ. The *Streptococcus pneumoniae* adhesin PsrP binds to keratin 10 on lung cells. *Mol Microbiol* 2009; 73: 663–679.
7. Pracht D, Elm C, Gerber J *et al.* PavA of *Streptococcus pneumoniae* modulates adherence, invasion, and meningeal inflammation. *Infect Immun* 2005; 73: 2680–2689.
8. Nelson AL, Ries J, Bagnoli F *et al.* RrgA is a pilus-associated adhesin in *Streptococcus pneumoniae*. *Mol Microbiol* 2007; 66: 329–340.
9. Barocchi MA, Ries J, Zogaj X *et al.* A pneumococcal pilus influences virulence and host inflammatory responses. *Proc Natl Acad Sci USA* 2006; 103: 2857–2862.
10. Hava DL, Camilli A. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Mol Microbiol* 2002; 45: 1389–1406.
11. LeMieux J, Hava DL, Basset A, Camilli A. RrgA and RrgB are components of a multisubunit pilus encoded by the *Streptococcus pneumoniae* rlrA pathogenicity islet. *Infect Immun* 2006; 74: 2453–2456.
12. Gianfaldoni C, Censini S, Hilleringmann M *et al. Streptococcus pneumoniae* pilus subunits protect mice against lethal challenge. *Infect Immun* 2007; 75: 1059–1062.
13. Basset A, Trzcinski K, Hermos C *et al.* Association of the pneumococcal pilus with certain capsular serotypes but not with increased virulence. *J Clin Microbiol* 2007; 45: 1684–1689.
14. Moschioni M, Donati C, Muzzi A *et al. Streptococcus pneumoniae* contains 3 rlrA pilus variants that are clonally related. *J Infect Dis* 2008; 197: 888–896.
15. Aguiar SI, Serrano I, Pinto FR, Melo-Cristino J, Ramirez M. The presence of the pilus locus is a clonal property among pneumococcal invasive isolates. *BMC Microbiol* 2008; 8: 41–51.
16. Moschioni M, De Angelis G, Melchiorre S *et al.* Prevalence of pilus encoding islets among acute otitis media *Streptococcus pneumoniae* isolates from Israel. *Clin Microbiol Infect* 2010; 16: 1501–1507.
17. Imai S, Ito Y, Ishida T *et al.* Distribution and clonal relationship of cell surface virulence genes among *Streptococcus pneumoniae* isolates in Japan. *Clin Microbiol Infect* 2011; 17: 1409–1414.
18. Donati C, Hiller NL, Tettelin H *et al.* Structure and dynamics of the pan-genome of *Streptococcus pneumoniae* and closely related species. *Genome Biol* 2010; 11: R107–R126.
19. Black S. The volatile nature of pneumococcal serotype epidemiology: potential for misinterpretation. *Pediatr Infect Dis J* 2010; 29: 301–303.
20. O'Brien KL, Nohynek H. Report from a WHO working group: standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae*. *Pediatr Infect Dis J* 2003; 22: 133–140.
21. Turner P, Hinds J, Turner C *et al.* Improved detection of nasopharyngeal co-colonization by multiple pneumococcal serotypes using latex agglutination or molecular serotyping by microarray. *J Clin Microbiol* 2011; 49: 1784–1789.
22. Adegbola RA, Hill PC, Secka O *et al.* Serotype and antimicrobial susceptibility patterns of isolates of *Streptococcus pneumoniae* causing invasive disease in The Gambia 1996–2003. *Trop Med Int Health* 2006; 11: 1128–1135.
23. Park IH, Park S, Hollingshead SK, Nahm MH. Genetic basis for the new pneumococcal serotype, 6C. *Infect Immun* 2007; 75: 4482–4489.
24. Enright MC, Spratt BG. A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. *Microbiology* 1998; 144: 3049–3060.
25. Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J Bacteriol* 2004; 186: 1518–1530.
26. Melegaro A, Gay NJ, Medley GF. Estimating the transmission parameters of pneumococcal carriage in households. *Epidemiol Infect* 2004; 132: 433–441.
27. Reis JN, Palma T, Ribeiro GS *et al.* Transmission of *Streptococcus pneumoniae* in an urban slum community. *J Infect* 2008; 57: 204–213.

Microbiology Papers in Press. Published October 27, 2011 as doi:10.1099/mic.0.054270-0

Point mutations in *wchA* are responsible for non-typeability of two invasive *Streptococcus pneumoniae* isolates.

Running title: *wchA* mutations in non-typeable *S. pneumoniae.*

Contents Category: Cell and molecular biology of microbes.

Sara Melchiorre[1*], Romina Camilli[2*], Agostina Pietrantoni[3], Monica Moschioni[1], Francesco Berti[1], Maria Del Grosso[2], Fabiana Superti[3], Michèle Anne Barocchi[1], Annalisa Pantosti[2**]

[1]Novartis Vaccines & Diagnostics, Siena, [2]Department of Infectious, Parasitic and Immune-mediated Diseases, and [3]Department of Technology and Health, Istituto Superiore di Sanità, Rome, Italy

* Equally contributed to this work

**Corresponding author

Mailing address:

Dr. Annalisa Pantosti

Department of Infectious, Parasitic and Immune-mediated Diseases

Istituto Superiore di Sanità

Viale Regina Elena 299, 00161 Rome, Italy.

Phone: +39 0649902852

Fax: +39 0649902886

E-mail: annalisa.pantosti@iss.it.

Summary: 209 words

Main text: 2857

Figures: 3

Tables: 1 in the supplementary materials

SUMMARY

Non-typeable *Streptococcus pneumoniae* (NTPn) strains are typically isolated from nasopharyngeal carriage or from conjunctivitis. Since the isolation of NTPn from invasive disease is rare, we characterized the genetic basis for non-typeability of two isolates obtained in Italy from two cases of bacteremic pneumonia. Multi Locus Sequence Typing (MLST) revealed that both NTPn belonged to ST191, which, according to the MLST database, is associated with serotype 7F. Sequencing of the capsular locus (*cps*) confirmed the presence of a 7F *cps* in both strains and revealed the existence of distinct single point mutations in the *wchA* gene (a glycosyltransferase), both leading to the translation of proteins truncated at the C-terminus. To verify if these mutations were responsible for non-typeability of the isolates, a functional 7F WchA was over-expressed in both NTPn. The two NTPn along with their WchA over-expressing derivatives were analyzed by Transmission Electron Microscopy and by high-resolution magic angle spinning NMR spectroscopy. Both NTPn were devoid of a polysaccharide capsule and WchA over-expression was sufficient to restore the assembly of a serotype 7F capsule on the surface of the two NTPn. In conclusion, we identified two new naturally-occurring point-mutations leading to the non-typeability in pneumococcus and demonstrated that WchA is essential for the biosynthesis of the serotype 7F capsule.

INTRODUCTION

The Gram-positive pathogen *Streptococcus pneumoniae* is a major cause of community acquired pneumonia as well as of upper respiratory tract infections such as acute otitis media and sinusitis, and invasive diseases like meningitis, bacteremia, and endocarditis. However, *S. pneumoniae* is also a commensal of the upper respiratory tract, especially of young children, which represent the reservoir for pneumococcal transmission within the community.

The ability of the pneumococcus to invade sterile sites is related to the expression of a polysaccharide capsule, which provides a barrier against host-cell mediated phagocytosis and allows bacterial persistence in the blood (Brown *et al.*, 1983). In addition, capsular polysaccharides (CPSs) are immunogenic and antibodies against CPSs provide protection against pneumococcal diseases; therefore, all of the current pneumococcal vaccines are based on a combination of different CPSs, either unconjugated or conjugated to a carrier protein (Gladstone *et al.*, 2011).

To date, 93 capsular types (serotypes) have been described on the basis of their different genetic, biochemical and antigenic properties (Bentley *et al.*, 2006; Calix & Nahm, 2010; Park *et al.*, 2007). Each pneumococcal serotype corresponds to a distinct CPS, synthesized on the bacterial surface by enzymes which are encoded within a single chromosomal locus (*cps*) with few remarkable exceptions (Llull *et al.*, 1999). The sequencing of different *cps* loci revealed the presence of a core of genes common to all capsular types, namely *wzg*, *wzh*, *wzd*, and *wze* (also known as *cpsA, B, C and D*), and of additional genes that are CPS-specific (Bentley *et al.*, 2006).

The different capsular types can be distinguished in the laboratory by reaction with type-specific antisera, evidenced by the capsular swelling or Quellung reaction (Austrian, 1976). Only a minority of *S. pneumoniae* strains does not react with the anticapsular antisera; these strains, defined as non-typeable pneumococci (NTPn), are thought to be non-encapsulated. NTPn are tipically isolated from carriage (Andrade *et al.*, 2010; Sa-Leao *et al.*, 2006), but are also found in conjunctivitis

74 (Carvalho *et al.*, 2003), acute otitis media (Xu *et al.*, 2011) and, rarely, in invasive diseases (Beall
75 *et al.*, 2006). Non-typeability can be due to different genetic modifications such as partial or
76 complete loss of the *cps* gene cluster (Hathaway *et al.*, 2004), presence of a novel gene (Baldry *et*
77 *al.*, 2009), single point mutations (Arrecubieta *et al.*, 1994) or sequence duplications (Waite *et al.*,
78 2001).
79 The characterization of NTPn by multilocus sequence typing (MLST), revealed the presence of
80 specific NT pneumococcal lineages which had lost the *cps* (Hanage *et al.*, 2006; Hathaway *et al.*,
81 2004) as well as of NTPn that were genetically related to encapsulated strains (Andrade *et al.*, 2010;
82 Hathaway *et al.*, 2004).
83 In this report, we investigated the genetic basis for non-typeability of two NTPn isolates responsible
84 for invasive disease. The strains were isolated in the same Italian hospital during 2006 through the
85 surveillance network of invasive pneumococcal disease (Gherardi *et al.*, 2009) from two cases of
86 bacteremic pneumonia.
87

87 # METHODS

88

89 **Bacterial strains and growth conditions.** *S. pneumoniae* AP422 and AP426 were isolated from
90 the blood of two adult patients with pneumonia admitted to the same Italian hospital in 2006. *S.*
91 *pneumoniae* AP425 (clinical isolate, serotype 7F) was used as a control. Strains were grown at
92 37°C in 5% CO_2-enriched atmosphere on Columbia agar plates (Oxoid, Cambridge, UK) with 5%
93 sheep blood or on Tryptic Soy Agar plates (Becton Dickinson, Cockeysville, MA) supplemented
94 with 10 mg/l colistine, 5 mg/l oxolinic acid and 5% sheep blood. For genomic DNA extraction and
95 NMR spectroscopy, liquid cultures were grown statically at 37°C in 5% CO_2 in Todd Hewitt Broth
96 (Becton Dickinson, Cockeysville, MA) supplemented with 0.5% (w/w) yeast extract (THYE) until
97 an $A_{600} = 0.25$.

98

99 **Serotyping.** Serotyping was performed by means of the Latex agglutination test and the Quellung
00 reaction using antisera produced by the Statens Serum Institute (Copenhagen, Denmark).

01

02 **Multi Locus Sequence Typing (MLST).** Genomic DNA from pneumococcal strains was prepared
03 with the Wizard Genomic DNA purification kit (Promega, Madison, WI, USA) following the
04 manufacturer's instructions. MLST was performed as previously described (Bagnoli *et al.*, 2008).

05

06 **7F *cps* detection and sequencing.** The amplification and sequencing of 7F *cps* was performed on
07 the genomic DNA of strains AP422, AP426 and AP425 using the primers listed in Table S1,
08 generating short overlapping DNA fragments. Amplicons were purified with magnetic carboxilate
09 beads (Agentcourt, Bioscience, Beverly, MA, USA); sequencing was carried out by use of an ABI
10 3730xl DNA Analyzer (Applied Biosystems, Carlsbad, CA, USA). Traces were visually inspected,

edited, assembled and analyzed with Vector NTI Advance 10.3.1 (Life technologies, Carlsbad, CA, USA).

Cloning and *S. pneumoniae* transformation. The shuttle plasmid pMU1328_*Pc_wchA* was obtained as described below. Briefly, the *wchA* gene of the AP425 strain and the constitutive promoter of erythromycin (*Pc*) (Gentile *et al.*, 2011) were amplified with the primers *wchA*_for GTGCGTGGATCCATGGATGAAAAAGGATTGAAAATT, *wchA*_rev CAGCGTGGATCCTCACTTCGCCCCTTCTCTCATAAA, and *Pc*_for GTGCGTGAATTCGAAACAGCAAAGAATGGCGGAAAC, *Pc*_rev CAGCGTGGATCCGTAATCACTCCTTCTTAATTACAA, respectively. The two PCR products, digested with BamHI and EcoRI-BamHI restriction enzymes, respectively, were cloned into pMU1328, containing an erythromycin resistance marker (Achen *et al.*, 1986). The ligations mixtures were then transformed into competent *Escherichia coli* DH10B and transformants were selected on plates supplemented with erythromycin (100 μg/ml). *Pc_wchA* insertion was confirmed by sequencing. pMU1328_*Pc_wchA* was then transformed in AP422 and AP426 by using conventional methods (Alloing *et al.*, 1998) and transformants were selected using erythromycin (1μg/ml), analyzed by PCR to confirm the presence of the plasmid, and further investigated for the presence of the capsule.

Transmission Electron Microscopy. For ultrastructural analysis of the polysaccharide capsule, bacterial strains were grown on Columbia agar plates and were prepared following the immune-stabilization method (Jacques & Gottschalk, 1997). Briefly, bacteria were resuspended in 0.5 ml of PBS pH 7.4 and incubated with 0.1 ml of type 7 antiserum (Statens Serum Institute) for 1 hr at 4°C. After immune-stabilization, bacterial cells were centrifuged at 2500 g for 10 minutes at 4°C and fixed in fixation solution containing 2.5 % (v/v) glutaraldehyde, 0.075 M lysine acetate in

cacodylate buffer (0.1 M, pH 7.0) with 0.075% (w/v) ruthenium red (CR buffer) for 2 hrs at 4°C. Controls processed using an unrelated antiserum (type 21), or without antiserum stabilization, were performed simultaneously. All bacterial samples were then washed three times with CR buffer and post-fixed with 1% (w/v) osmium tetroxide in CR buffer for 1 h at RT. Subsequently, samples were washed three times with cacodylate buffer, dehydrated through a graded series of ethanol and embedded in Agar 100 epoxy resin. Ultrathin sections of bacterial samples were placed on formvar-coated copper grids, post-stained with uranyl acetate and lead citrate, and examined with a Philips 208s electron microscope at 80 kV.

Type 7F capsular polysaccharide detection by high-resolution magic angle spinning NMR spectroscopy (HR-MAS NMR). Bacteria recovered from THYE were inactivated by 1% v/v formaldehyde treatment and then washed three times with PBS in deuterium oxide (D_2O, Sigma-Aldrich, St. Louis, MO, USA). Approximately 50 µl of compact pellet were inserted in a Kel-F disposable insert for 50 µl volume and then in a 4 mm MAS ZrO_2 rotor (Bruker, Madison, WI, USA). Proton HR-MAS NMR experiments were recorded by a Bruker Avance III 400 MHz spectrometer using a Bruker 4-mm HR-MAS probe. The spectra were recorded at 4500 Hz spin rate and 25°C. The ^1H spectra were acquired with the combination of a diffusion filter pulse sequence with gradient pulses (diffusion filter 85%), to remove the low molecular weight species free in solution, and a Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence [90-(τ-180-τ)$_n$-acquisition] as T2 filter (76.8 ms), to remove the broad signals of larger molecular species. 5 mg of purified 7F polysaccharide (Merck & Co., Whitehouse Station, NJ, USA) were solubilized in 0.75 ml of D_2O (D_2O, Sigma-Aldrich, St. Louis, MO, USA), the solution inserted in 5-mm NMR tube (Wilmad, Vineland, NJ, USA) and proton NMR experiment recorded at 25°C by the same spectrometer, using a 5-mm broadband probe (Bruker, Madison, WI, USA). ^1H NMR spectrum was recorded using a standard one-pulse experiment. Both the solid- and liquid-state NMR spectra were collected with 32

k data points over a 10 ppm spectral width. The transmitter was set at the HDO frequency, which was also used as reference signal (4.79 ppm). The TOPSPIN 2.1 software package (Bruker, Madison, WI, USA) was used for data acquisition and processing of all spectra.

RESULTS

S. pneumoniae NT invasive strains AP422 and AP426 contain a 7F capsular locus with point mutations in *wchA.*

The two invasive *S. pneumoniae* strains AP422 and AP426 were defined NT since serotype testing (confirmed by the Statens Serum Institute, Copenhagen) did not reveal positivity for any capsular type. In order to verify if these strains belonged to a clonal lineage specifically referable to NT pneumococci, they were subjected to MLST analysis. Both AP422 and AP426 were ST191, which, according to a search performed on the *S. pneumoniae* MLST database (www.mlst.net, accessed on 08.18.2011), is mainly (96%) associated with serotype 7F. Using PCR amplification, 7F *cps* was detected in both AP422 and AP426 and then sequenced. Comparison with two reference 7F *cps* sequences (GenBank accession numbers CR931643 and ABFT01000009) revealed the presence of the entire 7F locus with an overall nucleotide identity of 99.8% in both strains. Sequence analysis highlighted the existence of a distinct single point mutation in the *wchA* gene in both strains. With respect to the reference *wchA* nucleotide sequence, AP422 *wchA* showed a frameshift mutation due to the deletion of an adenine in position 910, while AP426 *wchA* showed a G1081T transvertion, which creates a new stop codon. Both mutations lead to the formation of a shorter *wchA* gene associated with the translation of a protein truncated at the C-terminus (Fig. 1). The NCBI Conserved Domain algorithm (www.ncbi.nlm.nih.gov/structure/cdd/cdd.shtml) predicts the WchA catalytic domain to be at the C-terminus of the protein (Fig. 1), so that the truncated WchA proteins coded by AP422 and AP426 are predicted to be enzymatically inactive. In the polysaccharide biosynthesis pathway, WchA (also known as CpsE) is the initial glucose phosphate-transferase, responsible for the linkage of an activated glucose phosphate to the lipid carrier (Pelosi *et al.*, 2005). Since WchA has been demonstrated to be required for the biosynthesis of other capsular

88 types (Pelosi *et al.*, 2005), the mutations identified in *wchA* in both AP422 and AP426 are likely to
89 be responsible for their non-typeability.
90
91 **Over-expression of a functional WchA restores the ability of *S. pneumoniae* to synthesize the**
92 **7F capsule.**
93 With the aim to demonstrate that the mutations present in *wchA* were sufficient to account for the
94 lack of capsule in AP422 and AP426, a functional WchA was over-expressed in both NTPn strains.
95 AP422 and AP426 were transformed with the recombinant plasmid pMU1328_*Pc_wchA* to
96 generate the respective recombinant strains AP422∇pMU1328_*Pc_wchA* and
97 AP426∇pMU1328_*Pc_wchA*. The two transformed NTPn yielded a positive reaction with the 7F
98 antiserum in the Quellung reaction. To confirm that WchA over-expression had restored the ability
99 to synthesize a capsule at levels comparable to a wild type capsulated strain, the NTPn isolates
00 along with their WchA over-expressing derivatives were analyzed by TEM and by HR-MAS NMR
01 using the serotype 7F strain AP425 as a positive control. For TEM analysis, the two NTPn strains
02 along with AP425 and the WchA over-expressing derivatives were stabilized with type 7 antiserum
03 before processing for TEM observation. This method was chosen as an alternative of the standard
04 fixation protocol with lysine and ruthenium red (Kreikemeyer *et al.,* 2011); indeed, the latter
05 appeared inadequate to visualize the 7F capsule polysaccharide (data not shown) since it does not
06 contain anionic charged moieties (Moreau *et al.*, 1988). After immune-stabilization a well-
07 preserved dense and thick capsular material was surrounding the AP425 bacterial cells (Fig. 2a),
08 while non-typeable AP422 (Fig. 2b) and AP426 (Fig. 2c) did not exhibit any characteristic
09 polysaccharide layer on the surface. In WchA over-expressing AP422 and AP426 (Fig. 2d and 2e)
10 TEM examination revealed cells surrounded by a capsular polysaccharide structure comparable to
11 that of AP425, although the capsule of WchA over-expressing AP422 appeared heterogeneous and
12 in some bacterial cells slightly irregular and thinner. In order to confirm the data obtained by TEM,

proton HR-MAS NMR spectra directly recorded on inactivated cells in the heterogeneous phase was performed. ^{1}H NMR spectrum collected on purified polysaccharide in liquid state clearly revealed specific peaks corresponding to the 7F capsular polysaccharide, such as the anomeric protons at 5.0÷5.6 ppm and the C6 methyl group of rhamnose at 1.2÷1.4 ppm (Fig. 3a). Spectra analysis showed the absence of the 7F specific peaks in the two NTPn and in AP422 and AP426 transformed with an empty pMU1328 plasmid (Fig. 3b) while the above mentioned 7F specific peaks were present both in the capsulated AP425 and in the WchA over-expressing AP422 and AP426 (Fig. 3b). Although signals of other ^{1}H-NMR sensitive molecules expressed on the bacterial surface are partially overlapped with the capsular polysaccharide peaks, the anomeric proton at 5.6 ppm and the C6 methyl group of rhamnose at 1.2-1.4 ppm fall in spectral windows without other peaks and the assignment results certain.

DISCUSSION

It is commonly assumed that NTPn lack the polysaccharide capsule, the principal pneumococcal virulence factor that acts by protecting the bacteria during the early phase of infection. Capsule deficiency impairs the capability of NTPn to evade the host's immune defence so that only rarely these strains are able to persist in the blood and cause invasive disease. The mechanism by which these strains overcome complement activity and phagocytosis during infection remains largely unknown. Other pneumococcal structures, such as the surface proteins PspA (Ren *et al*., 2003) and PspC (Jarva *et al*., 2002), are able to inhibit complement activation and deposition and likely contribute to inhibition of phagocytosis. Indeed, very few cases of invasive diseases due to NTPn have been reported in large epidemiological studies (Beall *et al.*, 2006).

Hathaway *et al*. reported that in approximately 30% of naturally occurring NTPn the *cps* was present. Non-typeability was assumed to be due to mutations in the *cps* genes, but sequencing was not performed and the gene(s) involved remained unknown (Hathaway *et al.*, 2004).

In this study, we have investigated the molecular basis responsible for non-typeability of two NT strains isolated from cases of bacteremic pneumonia in the same hospital during the same year. Molecular analysis of the two invasive NTPn, revealed that these strains were genetically related, since both belonged to ST191 and carried the 7F *cps*. However, they presented distinct point mutations in the *wchA* gene that were responsible for their non-encapsulated phenotype. *wchA* is responsible for the first step of capsule biosynthesis, since it codes for a glycosyltransferase that catalyzes the transfer of the initial sugar to a lipid acceptor, thus initiating the synthesis of the capsule repeat units (Pelosi *et al.*, 2005). Both mutations, introducing a premature stop codon in *wchA*, are expected to produce a glycosyltransferase truncated in its catalytic domain (Pelosi *et al.*, 2005). Re-establishment of a functional *wchA* in both NTPn restored the 7F capsule production as

clearly demonstrated by serotyping, TEM and HR-MAS NMR. HR-MAS spectroscopy appeared as a powerful methodology to detect the bacterial capsule and to confirm its identity. In fact, although precise quantification of the polysaccharide capsule content could not be provided due to the variable number of cells used as input in the analysis, this method clearly gave a semi-quantitative estimation of the specific polysaccharide expressed on the bacterial surface.

Our report is the first describing naturally occurring *wchA* mutations responsible for the NT phenotype in pneumococci. The essentiality of *wchA* for capsular assembly has previously been shown in laboratory mutants of other *S. pneumoniae* serotypes. Insertion-duplication mutagenesis or introduction of a premature stop codon in *wchA* made pneumococci of serotype 19F, 14, 9N, 13 and 15B unable to produce the capsule (Guidolin *et al.*, 1994; Kolkman *et al.*, 1996; Kolkman *et al.*, 1998). Furthermore, sequence duplications within *wchA* were found to be responsible for capsular phase variation in sorbarods-generated pneumococci of serotypes 3, 8 and 37 (Waite *et al.*, 2001; Waite *et al.*, 2003).

This study identified two new point mutations, involving *wchA,* responsible for the unencapsulated phenotype in two NTPn isolated from invasive disease, providing evidence regarding the key role of *wchA* also in serotype 7F capsule biosynthesis process. This result suggests that single point mutations represent a mechanism generating NTPn in nature besides the mechanisms already described such as down-regulation of capsule synthesis or loss of the capsule biosynthetic locus (Hathaway *et al.*, 2004; Hanage *et al*., 2006). It is intriguing that two genetically related strains, isolated from invasive diseases in the same hospital in the same year, underwent distinct point mutations in *wchA*, thus affecting the production of a type 7F capsule. The success of such strains in causing invasive disease might be due to a protection from the immune response exerted by other virulence factors and/or an immune deficiency of the host. Unfortunately, no data concerning the immunological status of patients from whom AP422 and AP426 were isolated are available. Furthermore, it is difficult to argue if the mutations occurred prior or during infection,

since we do not have serial isolates from the same patient. Extensive analyses on NTPn collections will ascertain whether *wchA* is prone to point mutations, leading to the loss of capsule, also in other NTPn clinical strains and estimate how frequently the single point mutation mechanism is used by pneumococci to silence capsule expression.

ACKNOWLEDGMENTS

The pMU1328 shuttle plasmid was kindly provided by Intercell AG (Austria). This work was supported in part by a grant from the Italian Ministry of Health (CCM) “Sorveglianza delle malattie batteriche invasive”. SM, MM, FB and MAB declare a potential conflict of financial interest as employees of Novartis Vaccines and Diagnostics. The authors have no other competing financial interests.

REFERENCES

Achen, M. G., Davidson, B. E. & Hillier, A. J. (1986). Construction of plasmid vectors for the detection of streptococcal promoters. *Gene* **45**, 45-49.

Alloing, G., Martin, B., Granadel, C. & Claverys, J. P. (1998). Development of competence in *Streptococcus pneumoniae*: pheromone autoinduction and control of quorum sensing by the oligopeptide permease. *Mol Microbiol* **29**, 75-83.

Andrade, A. L., Franco, C. M., Lamaro-Cardoso, J., Andre, M. C., Oliveira, L. L., Kipnis, A., Rocha, C. G., Andrade, J. G., Alves, S. L. & other authors (2010). Non-typeable *Streptococcus pneumoniae* carriage isolates genetically similar to invasive and carriage isolates expressing capsular type 14 in Brazilian infants. *J Infect* **61**, 314-322.

Arrecubieta, C., Lopez, R. & Garcia, E. (1994). Molecular characterization of cap3A, a gene from the operon required for the synthesis of the capsule of *Streptococcus pneumoniae* type 3: sequencing of mutations responsible for the unencapsulated phenotype and localization of the capsular cluster on the pneumococcal chromosome. *J Bacteriol* **176**, 6375-6383.

Austrian, R. (1976). The quellung reaction, a neglected microbiologic technique. *Mt Sinai J Med* **43**, 699-709.

Bagnoli, F., Moschioni, M., Donati, C., Dimitrovska, V., Ferlenghi, I., Facciotti, C., Muzzi, A., Giusti, F., Emolo, C. & other authors (2008). A second pilus type in Streptococcus pneumoniae is prevalent in emerging serotypes and mediates adhesion to host cells. *J Bacteriol* **190**, 5480-5492.

Baldry, S. J., Hinds, J., Gould, K., Lambertsen, L., Hanage, W., Antonio, M., Turner, P., Hermans, P. & Bentley, S. (2009). Sequencing the cps locus of non-typeable *Streptococcus pneumoniae* isolates. In *9th Eur Meet Mol Biol Pneum (Europneumo)*. Bern, Switzerland.

Beall, B., McEllistrem, M. C., Gertz, R. E., Jr., Wedel, S., Boxrud, D. J., Gonzalez, A. L., Medina, M. J., Pai, R., Thompson, T. A. & other authors (2006). Pre- and postvaccination clonal compositions of invasive pneumococcal serotypes for isolates collected in the United States in 1999, 2001, and 2002. *J Clin Microbiol* **44**, 999-1017.

Bentley, S. D., Aanensen, D. M., Mavroidi, A., Saunders, D., Rabbinowitsch, E., Collins, M., Danohoe, K., Harris, D., Murphy, L. & other authors (2006). Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. In *PLoS Genet*, pp. E31.

Brown, E. J., Hosea, S. W. & Frank, M. M. (1983). The role of antibody and complement in the reticuloendothelial clearance of pneumococci from the bloodstream. *Rev Infect Dis* **5 Suppl 4**, S797-805.

Calix, J. J. & Nahm, M. (2010). A new pneumococcal serotype, 11E, has a variably inactivated *wcjE* gene. *J Infec Dis* **202**.

Carvalho, M. G., Steigerwalt, A. G., Thompson, T., Jackson, D. & Facklam, R. R. (2003). Confirmation of nontypeable *Streptococcus pneumoniae*-like organisms isolated from outbreaks of epidemic conjunctivitis as *Streptococcus pneumoniae*. *J Clin Microbiol* **41**, 4415-4417.

Gentile, M. A., Melchiorre, S., Emolo, C., Moschioni, M., Gianfaldoni, C., Pancotto, L., Ferlenghi, I., Scarselli, M., Pansegrau, W. & other authors (2011). Structural and functional characterization of the *Streptococcus pneumoniae* RrgB pilus backbone D1 domain. *J Biol Chem* **286**, 14588-14597.

Gherardi, G., D'Ambrosio, F., Monaco, M., Camilli, R., De Florio, L., D'Ancona, F., Dicuonzo, G. & Pantosti, A. (2009). Population structure of invasive *Streptococcus pneumoniae* isolates in Italy prior to the implementation of the 7-valent conjugate vaccine (1999-2003). *Eur J Clin Microbiol Infect Dis* **28**, 99-103.

Gladstone, R. A., Jefferies, J. M., Faust, S. N. & Clarke, S. C. (2011). Continued control of pneumococcal disease in the UK - the impact of vaccination. *J Med Microbiol* **60**, 1-8.

Guidolin, A., Morona, J. K., Morona, R., Hansman, D. & Paton, J. C. (1994). Nucleotide sequence analysis of genes essential for capsular polysaccharide biosynthesis in *Streptococcus pneumoniae* type 19F. *Infect Immun* **62**, 5384-5396.

Hanage, W. P., Kaijalainen, T., Saukkoriipi, A., Rickcord, J. L. & Spratt, B. G. (2006). A successful, diverse disease-associated lineage of nontypeable pneumococci that has lost the capsular biosynthesis locus. *J Clin Microbiol* **44**, 743-749.

Hathaway, L. J., Stutzmann Meier, P., Battig, P., Aebi, S. & Muhlemann, K. (2004). A homologue of *aliB* is found in the capsule region of nonencapsulated *Streptococcus pneumoniae*. *J Bacteriol* **186**, 3721-3729.

Jacques, M. & Gottschalk, M. (1997). Use of monoclonal antibodies to visualize capsulare material of bacterial pathogens by conventional electron microscopy. *Microsc Microanal* **3**, 234-238.

Jarva, H., Janulczyk, R., Hellwage, J., Zipfel, P.F., Bjorck, L., Meri, S. (2002). *Streptococcus pneumoniae* evades complement attack and opsonophagocytosis by expressing the *pspC* locus-encoded Hic protein that binds to short consensus repeats 8-11 of factor H. J Immunol 168, 1886-1894.

Kolkman, M. A., Morrison, D. A., Van Der Zeijst, B. A. & Nuijten, P. J. (1996). The capsule polysaccharide synthesis locus of streptococcus pneumoniae serotype 14: Identification of the glycosyl transferase gene cps14E. *J Bacteriol* **178**, 3736-3741.

Kolkman, M. A., van der Zeijst, B. A. & Nuijten, P. J. (1998). Diversity of capsular polysaccharide synthesis gene clusters in *Streptococcus pneumoniae*. *J Biochem* **123**, 937-945.

Kreikemeyer, B., Gamez, G., Margarit, I., Giard, J. C., Hammerschmidt, S., Hartke, A. & Podbielski, A. (2011) Genomic organization, structure, regulation and pathogenic role of pilus constituents in major pathogenic Streptococci and Enterococci. *Int J Med Microbiol* **301**, 240-251.

Llull, D., Munoz, R., Lopez, R. & Garcia, E. (1999). A single gene (*tts*) located outside the cap locus directs the formation of *Streptococcus pneumoniae* type 37 capsular polysaccharide. Type 37 pneumococci are natural, genetically binary strains. *J Exp Med* **190**, 241-251.

Moreau, M., Richards, J. C., Perry, M. B. & Kniskern, P. J. (1988). Application of high-resolution n.m.r. spectroscopy to the elucidation of the structure of the specific capsular polysaccharide of *Streptococcus pneumoniae* type 7F. *Carbohydr Res* **182**, 79-99.

Park, I. H., Pritchard, D. G., Cartee, R., Brandao, A., Brandileone, M. C. & Nahm, M. H. (2007). Discovery of a new capsular serotype (6C) within serogroup 6 of *Streptococcus pneumoniae*. *J Clin Microbiol* **45**, 1225-1233.

Pelosi, L., Boumedienne, M., Saksouk, N., Geiselmann, J. & Geremia, R. A. (2005). The glucosyl-1-phosphate transferase WchA (Cap8E) primes the capsular polysaccharide repeat unit biosynthesis of *Streptococcus pneumoniae* serotype 8. *Biochem Biophys Res Commun* **327**, 857-865.

Ren, B., Szalai, A. J., Thomas, O., Hollingshead, S. K., Briles, D.E.(2003). Both family 1 and family 2 PspA proteins can inhibit complement deposition and confer virulence to a capsular serotype 3 strain of *Streptococcus pneumoniae*. *Infect Immun* **71**, 75-85.

Sa-Leao, R., Simoes, A. S., Nunes, S., Sousa, N. G., Frazao, N. & de Lencastre, H. (2006). Identification, prevalence and population structure of non-typable *Streptococcus pneumoniae* in carriage samples isolated from preschoolers attending day-care centres. *Microbiology* **152**, 367-376.

Waite, R. D., Struthers, J. K. & Dowson, C. G. (2001). Spontaneous sequence duplication within an open reading frame of the pneumococcal type 3 capsule locus causes high-frequency phase variation. *Mol Microbiol* **42**, 1223-1232.

Waite, R. D., Penfold, D. W., Struthers, J. K. & Dowson, C. G. (2003). Spontaneous sequence duplications within capsule genes cap8E and tts control phase variation in *Streptococcus pneumoniae* serotypes 8 and 37. *Microbiology* **149**, 497-504.

Xu, Q., Kaur, R., Casey, J. R., Sabharwal, V., Pelton, S. & Pichichero, M. E. (2011). Nontypeable *Streptococcus pneumoniae* as an otopathogen. *Diagn Microbiol Infect Dis* **69**, 200-204.

FIGURE LEGENDS

Fig. 1. Alignment of the WchA amino acids sequences of the reference strain 554/62 (GenBank accession number CR931643) and the two NTPn AP422 and AP426. In both AP422 and AP426 *wchA* is translated as a C-terminal truncated protein. Bold types indicate the catalytic domain as predicted by NCBI Conserved Domain; italic types indicate an alternative C-terminus due to a frame shift in the nucleotide sequence of AP422.

Fig. 2. Transmission Electron Microscopy (TEM) of *S. pneumoniae* strains incubated with type 7 antiserum and stained with ruthenium red. (a): AP425 (serotype 7F); (b): AP422 (NT); (c): AP426 (NT); (d): WchA over-expressing AP422 strain; (e): WchA over-expressing AP426 strain. Over-expression of WchA restored capsule production in both AP422 and AP426.

Fig. 3. Proton NMR spectra of 7F purified polysaccharide and *S. pneumoniae* strains. (a) Schematic representation of the chemical structure of 7F polysaccharide. (b) Proton HR-MAS NMR spectra recorded in heterogeneous phase on inactivated bacterial cells. The presence or the absence of the capsule on the bacterial surface, based on spectra analysis, is indicated for each strain. Red arrows indicate anomeric proton at 5.6 ppm, while blue arrows indicate the C6 methyl group of rhamnose at 1.2÷1.4 ppm. 7F purified polysaccharide solubilised in denaturated water is used as a positive control.

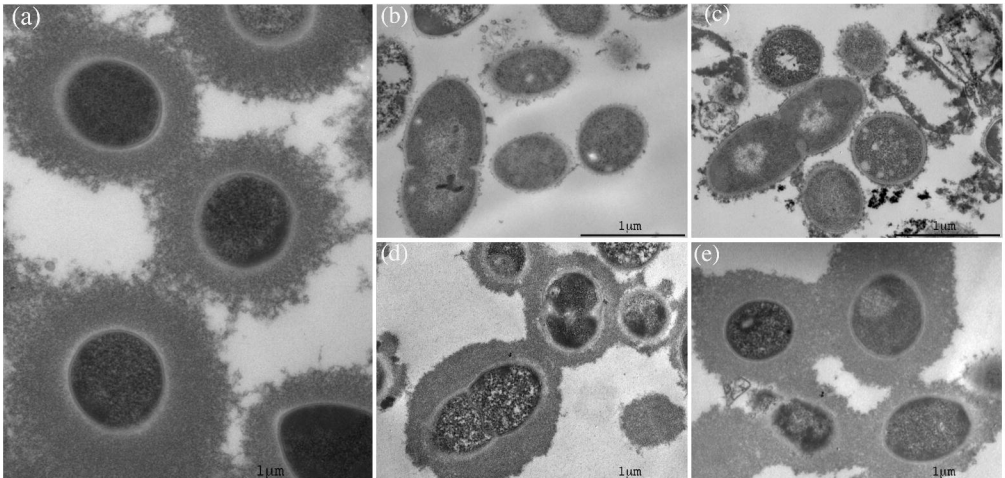
554/62 1 MDEKGLKIFLAVLQSIIVILLVYFLSFVREAELERSSMVILYLLHFFVFYFSSYGNNFFKRGYLVEFNSTIRYIFFFAIAISVLNFFIAERFSISRRGMV
AP422 1 MDEKGLKIFLAVLQSIIVILLVYFLSFVREAELERSSMVILYLLHFFVFYFSSYGNNFFKRGYLVEFNSTIRYIFFFAIAISVLNFFIAERFSISRRGMV
AP426 1 MDEKGLKIFLAVLQSIIVILLVYFLSFVREAELERSSMVILYLLHFFVFYFSSYGNNFFKRGYLVEFNSTIRYIFFFAIAISVLNFFIAERFSISRRGMV

554/62 101 YFLTLEGISLYLLNFLVKKYWKHVFFNLKNSKKILLLTVTKNMEKVLDKLLESDELSWKLVAVSVLDKSDFQHDKIPVIEKEKIIEFATHEVVDEVFVNL
AP422 101 YFLTLEGISLYLLNFLVKKYWKHVFFNLKNSKKILLLTVTKNMEKVLDKLLESDELSWKLVAVSVLDKSDFQHDKIPVIEKEKIIEFATHEVVDEVFVNL
AP426 101 YFLTLEGISLYLLNFLVKKYWKHVFFNLKNSKKILLLTVTKNMEKVLDKLLESDELSWKLVAVSVLDKSDFQHDKIPVIEKEKIIEFATHEVVDEVFVNL

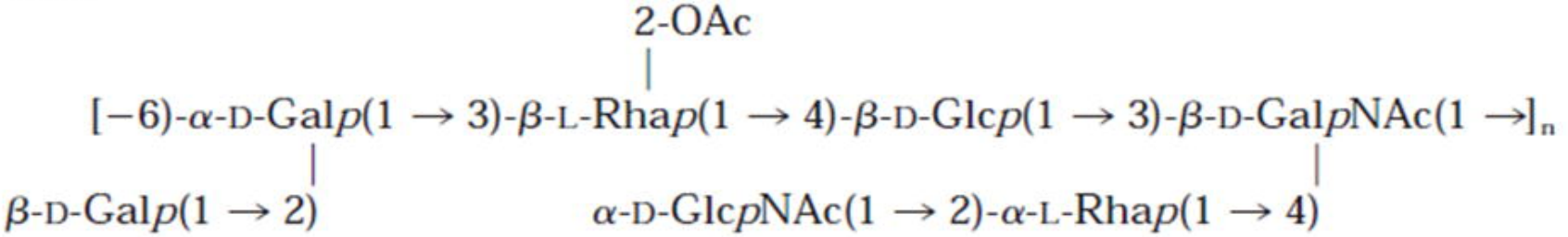
554/62 201 PGESYDIGEIISRFETMGIDVTVNLKAFDKNLGRNKQIYEMVGLNVVTFSTNFYKTSHVIS**KRILDICGATIGLILFAIASLVLVPLIRKDGGPAIFAQT**
AP422 201 PGESYDIGEIISRFETMGIDVTVNLKAFDKNLGRNKQIYEMVGLNVVTFSTNFYKTSHVIS**KRILDICGATIGLILFAIASLVLVPLIRKDGGPAIFAQT**
AP426 201 PGESYDIGEIISRFETMGIDVTVNLKAFDKNLGRNKQIYEMVGLNVVTFSTNFYKTSHVIS**KRILDICGATIGLILFAIASLVLVPLIRKDGGPAIFAQT**

554/62 301 **RIGKNGRHFTFYKFRSMRSDAEAIKEQLMDQNTMQGGMFKIDNDPRVTKIGRFIRKTSLDELPQFWNVFIGDMSLVGTRPPTVDEYVQYTPEQKRRLSFK**
AP422 301 **RIGKM***VDILPFINSVRCGAMLKLSKNS*
AP426 301 **RIGKNGRHFTFYKFRSMRSDAEAIKEQLMDQNTMQGGMFKIDNDPRVTKIGRFIRKTSLD**

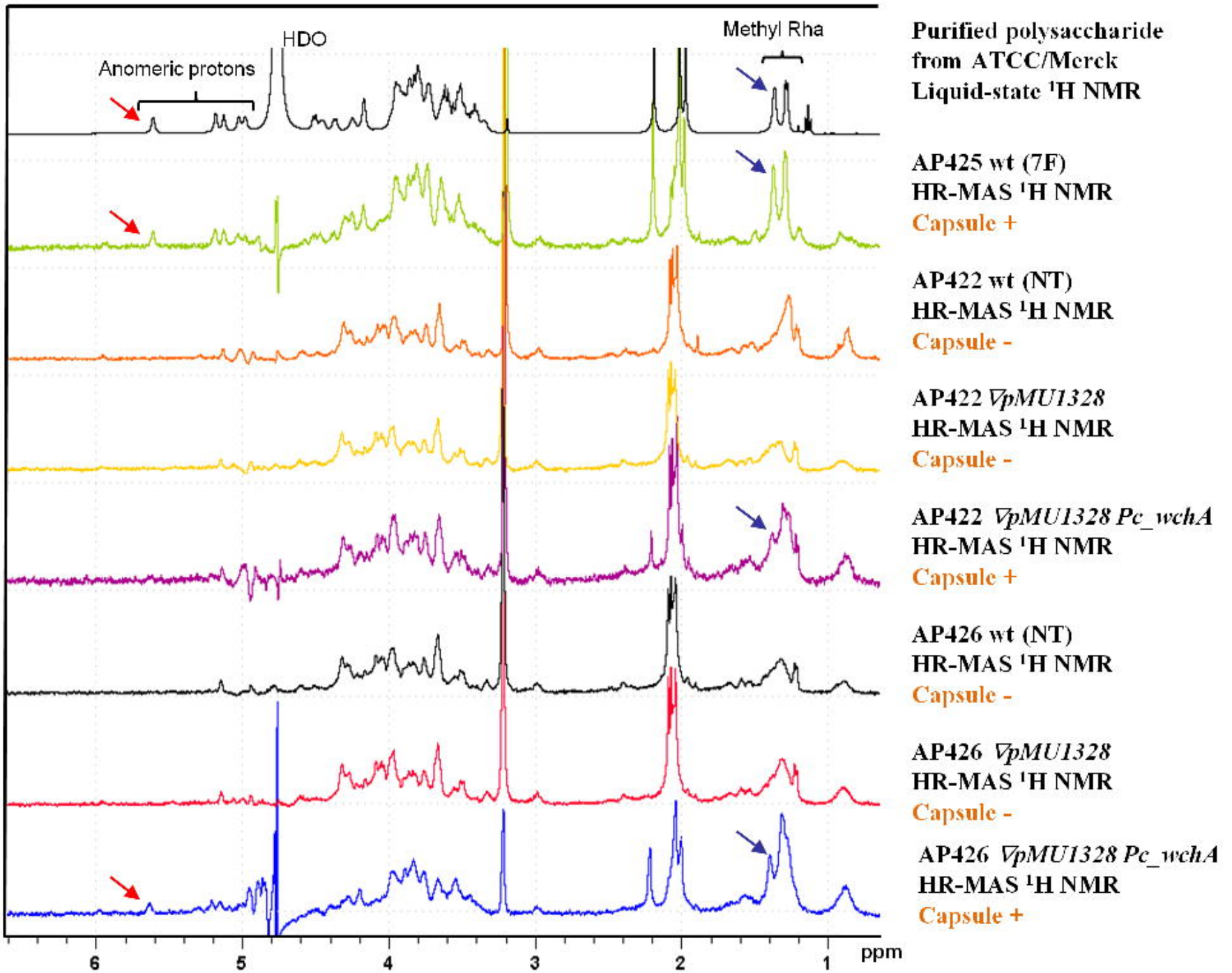
554/62 401 **PGITGLWQVSGRSKITDFDDVVKLDVAYIDNWTIWKDIEILLKTVKVVFMREGAK**
AP422
AP426


(a)
1 μm
(b)
1 μm
(c)
1 μm
(d)
1 μm
(e)
1 μm

(a)


2-OAc
[−6)-α-D-Galp(1 → 3)-β-L-Rhap(1 → 4)-β-D-Glcp(1 → 3)-β-D-GalpNAc(1 →]n
β-D-Galp(1 → 2)
α-D-GlcpNAc(1 → 2)-α-L-Rhap(1 → 4)

(b)


Anomeric protons
HDO
Methyl Rha
Purified polysaccharide from ATCC/Merck Liquid-state ¹H NMR
AP425 wt (7F) HR-MAS ¹H NMR Capsule +
AP422 wt (NT) HR-MAS ¹H NMR Capsule -
AP422 ∇pMU1328 HR-MAS ¹H NMR Capsule -
AP422 ∇pMU1328 Pc_wchA HR-MAS ¹H NMR Capsule +
AP426 wt (NT) HR-MAS ¹H NMR Capsule -
AP426 ∇pMU1328 HR-MAS ¹H NMR Capsule -
AP426 ∇pMU1328 Pc_wchA HR-MAS ¹H NMR Capsule +
6
5
4
3
2
1
ppm

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 286, NO. 16, pp. 14588–14597, April 22, 2011

Structural and Functional Characterization of the *Streptococcus pneumoniae* RrgB Pilus Backbone D1 Domain*[S]

Received for publication, November 11, 2010, and in revised form, February 16, 2011 Published, JBC Papers in Press, March 2, 2011, DOI 10.1074/jbc.M110.202739

Maria Antonietta Gentile[‡], Sara Melchiorre[‡], Carla Emolo[‡], Monica Moschioni[‡], Claudia Gianfaldoni[‡], Laura Pancotto[‡], Ilaria Ferlenghi[‡], Maria Scarselli[‡], Werner Pansegrau[‡], Daniele Veggi[‡], Marcello Merola[‡§], Francesca Cantini[¶], Paolo Ruggiero[‡], Lucia Banci[¶1], and Vega Masignani[‡2]

From the [‡]Novartis Vaccines and Diagnostics Research Center, Via Fiorentina 1, Siena 53100, Italy, the [¶]Magnetic Resonance Center, Department of Chemistry, University of Florence, Via L. Sacconi 6, 50019 Sesto Fiorentino, Italy, and the [§]Università di Napoli Federico II, Monte Sant'Angelo, Via Cintia, Napoli 80126, Italy

Streptococcus pneumoniae **expresses on its surface adhesive pili, involved in bacterial attachment to epithelial cells and virulence. The pneumococcal pilus is composed of three proteins, RrgA, RrgB, and RrgC, each stabilized by intramolecular isopeptide bonds and covalently polymerized by means of intermolecular isopeptide bonds to form an extended fiber. RrgB is the pilus scaffold subunit and is protective *in vivo* in mouse models of sepsis and pneumonia, thus representing a potential vaccine candidate. The crystal structure of a major RrgB C-terminal portion featured an organization into three independently folded protein domains (D2–D4), whereas the N-terminal D1 domain (D1) remained unsolved. We have tested the four single recombinant RrgB domains in active and passive immunization studies and show that D1 is the most effective, providing a level of protection comparable with that of the full-length protein. To elucidate the structural features of D1, we solved the solution structure of the recombinant domain by NMR spectroscopy. The spectra analysis revealed that D1 has many flexible regions, does not contain any intramolecular isopeptide bond, and shares with the other domains an Ig-like fold. In addition, we demonstrated, by site-directed mutagenesis and complementation in *S. pneumoniae*, that the D1 domain contains the Lys residue (Lys-183) involved in the formation of the intermolecular isopeptide bonds and pilus polymerization. Finally, we present a model of the RrgB protein architecture along with the mapping of two surface-exposed linear epitopes recognized by protective antisera.**

Streptococcus pneumoniae is an important human pathogen responsible for diseases such as otitis media, pneumonia, sepsis, and meningitis (1–6). However, *S. pneumoniae* is also a common inhabitant of the respiratory tract of children and healthy adults. This carriage state could represent a risk factor for the development of respiratory diseases but also the source for pneumococcal transmission to other individuals (7–9). Like most streptococci, *S. pneumoniae* decorates its surface with long filaments known as pili (10–14). Pneumococcal pili have previously been associated with virulence and the capability of the microorganism to adhere better to epithelial cells and to colonize the nasopharynx (10, 15, 16). The pneumococcal pilus is a multimeric structure consisting of three proteins (RrgA, RrgB, and RrgC) polymerized by three sortase enzymes (SrtC1, SrtC2, and SrtC3) through the formation of covalent intermolecular isopeptide bonds (17–21). In particular, multiple copies of RrgB are polymerized to form the scaffold of the pilus, whereas the major adhesin, RrgA, and the putative anchor, RrgC, are localized at the tip and at the base of the pilus, respectively (15, 22, 23).

Recently, the structure of a major portion of RrgB (residues 184–627) was solved at a 1.6 Å resolution (24) and revealed an organization into three independently folded IgG-like domains (D2, D3, and D4, residues 184–326, 326–446, and 446–627, respectively). On the contrary, the structure of the RrgB N-terminal region (D1, residues 1–184), likely constituting a fourth independently folded domain, remained unsolved due to the failure to obtain the crystals of the full-length (FL)[3] RrgB (24). Interestingly, each of the D2, D3, and D4 domains is stabilized by one intramolecular isopeptide bond. These covalent linkages, formed between Lys and Asn residues, have been found in other pilus proteins (19, 25–28) and are thought to play a role similar to that of disulfide bonds; they confer in fact a rigid molecular architecture to the pili and make them less susceptible to proteolytic cleavage (Fig. 1).

In the pilus backbone assembly RrgB molecules are linked together by sortases through intermolecular isopeptide bond formation between a Thr within the C-terminal LP*X*TG motif of a molecule and a Lys located at the N terminus of the following molecule (18, 26, 29) (Fig. 1). In *Corynebacterium diphtheriae*, where the general principles of pilus assembly were first

* This work was supported by Ministero dell'Istruzione, dell'Università e della Ricerca (Fondi per gli Investimenti della Ricerca di Base-Proteomica RBRN07BMCT) and Integrated Structural Biology Infrastructure for Europe Contract 211252. Because some of the authors are employees of Novartis Vaccines, there are competing financial interests.

[S] The on-line version of this article (available at http://www.jbc.org) contains Materials and Methods, supplemental Tables S1–S6, and supplemental Figs. S1–S6.

The atomic coordinates and structure factors (code 2L40) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

The NMR chemical shifts have been deposited in the BioMagResBank (accession no. 17246).

[1] To whom correspondence may be addressed. Tel.: 39-055-4574263; Fax: 39-055-4574253; E-mail: banci@cerm.unifi.it.

[2] To whom correspondence may be addressed. Tel.: 39-0577-243319; Fax: 39-0577-243564; E-mail: vega.masignani@novartis.com.

[3] The abbreviations used are: FL, full-length; BisTris, bis(2-hydroxyethyl)iminotris(hydroxymethyl)methane; cryo-EM, cryoelectron microscopy; cfu, colony-forming units; HSQC, heteronuclear single quantum coherence; PDB, Protein Data Bank; r.m.s.d., root mean square deviation; WB, Western blot.

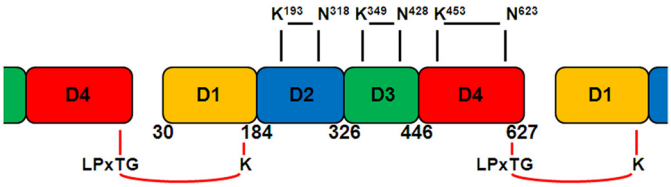


FIGURE 1. **Schematic representation of pilus backbone protein RrgB.** Pilus scaffold is composed by multiples copies of RrgB protein in a head-to-tail arrangement. Pilus polymerization occurs through intermolecular isopeptide bonds (*red*), whereas each RrgB protein is stabilized by intramolecular isopeptide bonds (*black*). Lys-183 as a residue involved in the intermolecular bond has been identified in the present work.

established, the functional Lys is located within a conserved YPKN "pilin" motif (18, 27, 30). Nevertheless, this sequence is not absolutely required for polymerization as demonstrated by studies on the Spy0128 pilin of *Streptococcus pyogenes*, where the lysine forming the intermolecular isopeptide bond and responsible for pilus polymerization is located into ^{159}GSKVPI^{164} motif even though the YPKN pilin motif is also present (26, 31).

RrgB, along with the other two pilus proteins RrgA and RrgC, was previously shown to confer protection in mouse models of infection and therefore is regarded as a potential candidate for a new generation of protein-based vaccines (32, 33). We have investigated the protective ability of the single recombinant D1, D2, D3, and D4 domains of RrgB in a mouse model of sepsis, and here we provide evidence that D1 is the most protective, followed by D4. Furthermore, we present the solution structure of the recombinant D1 obtained by NMR spectroscopy and show that Lys-183 of D1 is engaged in the intermolecular isopeptide bond formation during pilus polymerization. Finally, we propose a possible model of the entire RrgB molecule and show the positions of two linear epitopes possibly involved in the protection mechanism.

MATERIALS AND METHODS

Bacterial Strains and Cultures—For the animal experiments, the *S. pneumoniae* TIGR4 (serotype 4) strain was used. Bacteria were grown, frozen in aliquots at −80 °C, and titrated as already reported (32). Immediately prior to challenge, frozen aliquots were thawed and diluted in PBS to reach the working concentration.

Cloning and Protein Expression and Purification—Standard recombinant DNA techniques were used to construct the expression plasmids (pET21b$^+$; Novagen) and to express and purify the recombinant C-terminal His-tagged proteins (for details, see supplemental Materials and Methods, and the primers used are listed in supplemental Table S1). The affinity-purified proteins were subsequently used to immunize CD1 mice or rabbits for antibody generation (Charles River Laboratory) and BALB/c mice to evaluate the protective efficacy.

Complementation Plasmids—Wild-type or mutant *rrgB* genes were amplified from chromosomal DNA of TIGR4 strain by PCR by using the primers listed in supplemental Table S2; point mutations were introduced by overlap extension PCR by using specific primers (supplemental Table S2). PCR products were then cloned into the complementation plasmid pMU1328 between the BamHI and SalI restriction sites (34). Expression of RrgB or RrgB mutated forms was under control of the erythromycin constitutive promoter (Pc) which was amplified with the primers listed in supplemental Table S2 and cloned immediately upstream *rrgB* (EcoRI, BamHI). All plasmids were confirmed by sequencing.

Generation of rrgB Deletion Mutants and rrgB Complementants—A TIGR4 Δ*rrgB* isogenic mutant was generated by allelic exchange. Fragments of ~500 bp upstream and downstream the target gene were amplified by PCR (oligonucleotides are listed in supplemental Table S2) and spliced into a kanamycin resistance cassette by using overlap extension PCR; the PCR fragments were then cloned into pGEMt (Promega) and transformed in *S. pneumoniae* with conventional methods (35). To select the bacteria in which the target gene was replaced with the resistance cassette, bacteria were plated on blood-agar plates with kanamycin (500 μg/ml). The presence of the isogenic mutation was confirmed by PCR and Western blot (WB) analysis. To obtain RrgB-complemented mutants, pMU1328 plasmids containing WT *rrgB* or *rrgB* mutated forms were transformed into TIGR4 Δ*rrgB* with conventional methods. Transformants selection was performed by supplementing media with kanamycin (500 μg/ml) and erythromycin (1 μg/ml). The complemented mutants were then analyzed by PCR; expression of FL WT RrgB or RrgB mutants was detected by WB analysis of whole cell lysates.

SDS-PAGE and Western Blot Analysis—SDS-PAGE analysis was performed using NuPAGE™ 4–12% BisTris gradient gels (Invitrogen) according to the manufacturer's instructions. Hi-Mark™ prestained HMW (Invitrogen) served as protein standard. Gels were processed for WB analysis by using standard protocols. Mouse polyclonal antibodies raised against recombinant His-tagged RrgB were used at 1/3000 dilution. Secondary goat anti-mouse IgG alkaline phosphatase-conjugated antibodies (Promega) were used at 1/5000, and signals were developed by using Western Blue Stabilized Substrate for Alkaline Phosphatase (Promega).

Animal Experiments—Animal studies were done in compliance with the current law approved by the local Animal Ethics Committee and authorized by the Italian Ministry of Health. Female, 6-week-old, specific pathogen-free BALB/c mice (Charles River) received three intraperitoneal immunizations, 2 weeks apart. Each dose was composed of 20 μg of either the single RrgB domains or the FL RrgB, or of a combination of the four RrgB domains (D1+D2+D3+D4), 10 μg each, along with 400 μg of aluminum hydroxide as an adjuvant, in a final volume of 200 μl of PBS. Control animals received the same course of saline plus adjuvant. Ten days after the third immunization, samples of sera were obtained from each animal and pooled according to the immunization group to be used in passive serum transfer experiment. Two weeks after the third immunization, each mouse was challenged intraperitoneally with a mean dose of 1.6×10^2 cfu of TIGR4. Bacteremia was evaluated 24 h after challenge, and mortality course monitored for 10 days after challenge as already reported (32). The animals were euthanized when they exhibited defined humane end points that had been preestablished for the study in agreement with Novartis Animal Welfare Policies.

For passive protection experiments, 8-week-old mice were used. Fifteen minutes before TIGR4 intraperitoneal challenge (10^2 cfu/mouse), each mouse received intraperitoneally 50 μl of pooled mouse sera against recombinant D1 or D4, or of control sera obtained immunizing with adjuvant plus saline.

Statistical Analysis—Bacteremia and mortality course were analyzed by the Mann-Whitney *U* test. Survival rates were analyzed by Fisher's exact test. One-tailed or two-tailed tests were used to compare immunized groups with the control group or each other, respectively. Values of $p \leq 0.05$ were considered and referred to as significant. Values of $p \leq 0.1$ were referred to as a trend.

Flow Cytometry on Entire Bacteria—TIGR4 were grown in Todd-Hewitt yeast extract broth to an exponential phase (A_{600} = 0.25), fixed with 2% formaldehyde, and then stained with pooled mouse antisera raised against FL RrgB or RrgB domains at 1:400 dilution. Mouse IgG were detected with FITC-conjugated goat anti-mouse IgG (Jackson Laboratories) at 1:100 dilution, and bacterial staining was analyzed by using a FACS-Calibur cytometer (Becton Dickinson). Sera from mice immunized with PBS plus adjuvant were used as negative control.

ELISA—96-well MaxiSorp^{TM} flat-bottom plates (Nunc) were coated with 0.2 μg/well FL RrgB overnight at 4 °C. Plates were then washed three times with PBS/0.05% Tween 20 and saturated for 1 h at 37 °C with PBS/1% BSA. Following three washing steps with PBS/0,05% Tween 20, the plates were incubated for 2 h at 37 °C with serial dilutions of the pooled mouse sera. After another three washing steps, bound antigen-specific mouse IgGs were revealed with alkaline phosphatase-conjugated goat anti-mouse IgG (Sigma), followed by the phosphatase alkaline substrate *p*-nitrophenyl phosphate (Sigma). The intensity of color was quantified with an ELISA plate reader at A_{405}. The antibody titer was expressed as the log10 of the reciprocal of the serum dilution that gave A_{405} = 1.

PepScan Analysis—Arrayed peptides were synthesized *in situ* on glass fiber membranes. Membranes were conditioned by wetting with ethanol and washing three times for 5 min in TTBS (50 mM Tris-HCl, pH 7.0, 137 mM NaCl, 2.7 mM KCl, 0.05% Tween 20). After overnight blocking at 4 °C in MBS (2% dry milk in TTBS), membranes were incubated for 1.5 h at 37 °C with polyclonal antisera (1:3000 in MBS) followed by secondary goat anti-mouse IgG alkaline phosphatase-conjugated antibodies (1:5000 in MBS; Promega), and signals were developed by using Western Blue Stabilized Substrate for Alkaline Phosphatase. For image processing, membranes were scanned using an Epson V750 Pro scanner at 800 dpi, 48-bit color depth and with gamma 1.0 full linear response.

NMR Characterization of RrgB D1 Domain—Expression and purification of labeled D1 were carried out as described in the supplemental Materials and Methods. NMR spectra were acquired at 298 K on Avance 900, 800, 700, and 500 MHz Bruker spectrometers, all equipped with a triple resonance cryoprobe. The NMR experiments, used for the backbone and the aliphatic side chain resonances assignment recorded on $^{13}C/^{15}N$ and ^{15}N enriched samples or on unlabeled D1 samples, are summarized in supplemental Table S3. Backbone dihedral angle constraints were derived from ^{15}N, $^{13}C'$, $^{13}C\alpha$, $^{13}C\beta$, and Ha chemical shifts, using TALOS+ (36). Distance constraints for structure determination were obtained from ^{15}N-edited and ^{13}C-edited three-dimensional NOESY-HSQC. 3131 meaningful proton-proton distance restraints (supplemental Table S4), with 114 φ and 120 ψ backbone dihedral angles restraints were included in structure calculations. The exchangeability of the backbone amide hydrogen nuclei with solvent protons was investigated through a ^{1}H-^{15}N HSQC experiment performed on a protein sample dialyzed against deuterated buffer for 3 days. Hydrogen bond constraints for the slowly deuterium-exchanging amide protons of the β-strands were introduced at later steps of the structure calculations.

Structure calculations were performed through iterative cycles of CYANA-2.1 (37) followed by restrained energy minimization with the AMBER 10.0 Package in explicit water solvent (38). The quality of the structures was evaluated by using the iCING validation program (for details, see supplemental Table S4). The program MOLMOL was subsequently used for structure analysis (39).

The final bundle of 20 conformers of D1 has an average target function of 1.36 ± 0.13 (CYANA units). The average backbone r.m.s.d. value (over residues 28–183) is 0.71 ± 0.19Å, and the all-heavy atoms r.m.s.d. value is 0.96 ± 0.16. Per-residue r.m.s.d. values are shown in supplemental Fig. S1.

Heteronuclear Relaxation Data—The dynamic properties of D1 have been characterized experimentally through ^{15}N relaxation measurements. ^{15}N longitudinal and traverse relaxation rates (40) and $^{15}N\{^{1}H\}$ NOEs (41) were recorded at 298 K at 500 MHz, using a protein concentration of 0.8 mM.

The average backbone ^{15}N longitudinal R_1 and transversal R_2 relaxation rates and $^{15}N\{^{1}H\}$ NOE values are 1.45 ± 0.1 s^{-1}, 16.18 ± 1.5 s^{-1}, and 0.71 ± 0.04, respectively (supplemental Fig. S2). They are essentially homogeneous along the entire polypeptide sequence, with the exception of residues located at the C and N termini and three loop regions (56–69, 148–162, and 173–177). The correlation time for the molecule tumbling (τ_c), as estimated from the R_2/R_1 ratio, is 11.49 ± 1.9 ns, consistent with the molecular mass of the monomeric protein. The relaxation data were analyzed according to the model-free approach of Lipari and Szabo (42, 43) using TENSOR2 (44) (supplemental Fig. S2).

Rigid Body Fitting—The procedure used to accommodate the NMR D1 structure into the EM map of the whole pilus previously generated (24) followed the same procedures used for the fitting of the RrgBD2-D4 x-ray crystal structure and the D1 computer model into the EM pilus map (22). A preliminary rigid body fitting of the D2–D4 crystal fragment was performed by using CHIMERA (45–47) followed by a rigid body fitting of the D1 NMR coordinates into the leftover N-terminal apical volume of the pilus. The D1 NMR coordinates were first fitted manually using CHIMERA by placing as much of the atomic structure as possible into the EM density map, approximately in the position thought to be correct. This step was then followed by a rigid-body fitting using the "fit model in map" tool from CHIMERA. This tool calculates, for the selected atoms, a position that maximizes locally the sum of the densities. The evaluation of the correlation coefficient values, the resulting average map value at the fit atoms, the number of fits atoms outside

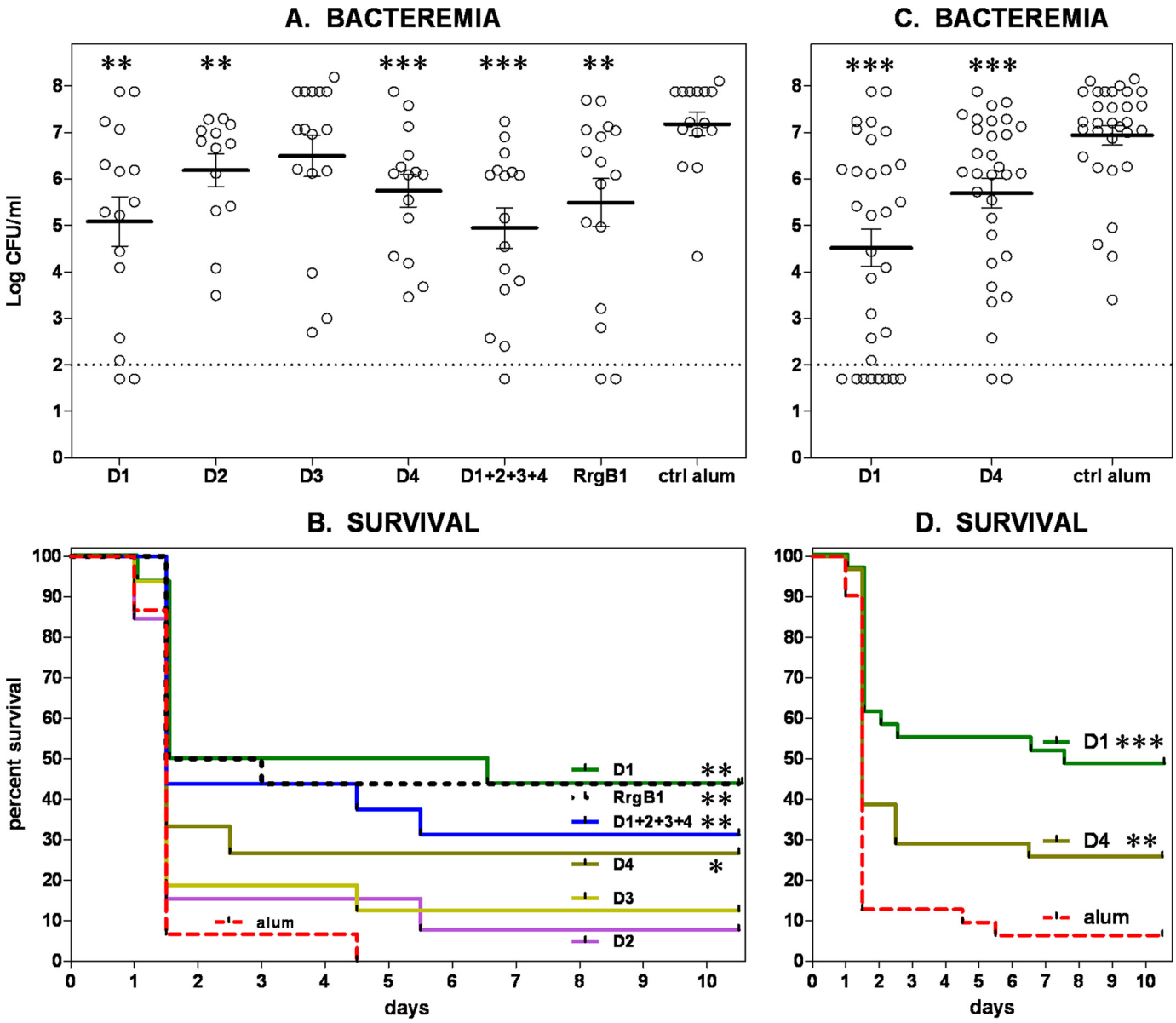


FIGURE 2. **RrgB domains are protective in active immunization experiments.** *A* and *C*, bacteremia. *Circles* represent the log cfu/ml of blood for individual animals; *horizontal bars* represent the mean value of the log cfu/ml ± S.E. for the group; the *dotted line* marks the detection limit (values under the *dotted line* correspond to animals in which no cfu were detected). The treatment groups are indicated on the *bottom*. *B* and *D*, survival. The survival course for each group is represented. The treatment groups are indicated close to each of the corresponding survival lines. *A* and *B*, $n = 13$ for D2 group, 15 for D4 and control (*alum ctrl*) groups, 16 for the remaining groups. *C* and *D*, $n = 31$ for each of the groups. ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

the lowest value contour surface displayed were the parameters used to assess the fit of the D1 molecule.

RESULTS

Distinct Domains of RrgB Provide Protection in Active Immunization Experiments—The protective efficacy of each of the four RrgB domains or of a mixture of them was assessed in a mouse model of intraperitoneal challenge (TIGR4) in two distinct experiments, performed under the same conditions, which were combined to reach $n = 13–16$ mice/group. The results are summarized in Fig. 2, *A* and *B*, and detailed statistical analysis is provided in supplemental Table S5*A*.

All RrgB domains except D3 afforded significant protection against bacteremia (Fig. 2*A*), giving a reduction of the cfu geometric mean by 1–2 logs with respect to the controls. These values were similar to those afforded by the FL RrgB and by the combination of the four domains D1+D2+D3+D4. The results of mortality are reported in Fig. 2*B* and supplemental Table S5*A*. D1 and D4 conferred significant increase of survival time, similar to FL RrgB and the combination D1+D2+D3+D4. In particular, the median survival for the D1 group was 2.5 days higher respect to that of the control group (4 *versus* 1.5 days). At the end of the mortality observation, a significant survival rate was found for D1 (44% survival), D4 (27%), FL RrgB (44%), and D1+D2+D3+D4 (31%) groups.

N-terminal D1 Domain Is the Most Protective in Vivo—Among the single RrgB domains, D1 and D4 showed the most significant protective efficacy and were therefore analyzed further in a larger group of mice. Four different experiments, carried out under the same conditions, were combined to

reach $n = 31$ animals/group. The results are shown in Fig. 2, *C* and *D*, and detailed statistical analysis is provided in supplemental Table S5*B*. In terms of bacteremia (Fig. 2*C*), both D1 and D4 afforded highly significant protection, with a cfu geometric mean by 2.6 and 1.5 logs lower, respectively, than that of the control group, and 8 animals from the D1 group in which cfu were undetectable. The reduction of bacteremia was significantly superior in D1 than in the D4 group.

In terms of mortality course (Fig. 2*D*), both D1 and D4 conferred significant protection. The increase of survival time afforded by D1 showed a better trend than that of D4. In particular, only for the D1 group was the median survival time higher than that of the control group (7.5 *versus* 1.5 days). At the end of the mortality observation, the D1 group showed the highest survival rate, *i.e.* 45% *versus* 21% observed for D4. An evident difference of survival rates between D1 and D4 groups was observed.

The possible relevance of antibodies in the protection elicited by D1 and D4 was investigated by a passive serum transfer experiment, with groups of 8 mice. The results are shown in Fig. 3, and detailed statistical analysis is provided in supplemental Table S5*C*. Both anti-D1 and anti-D4 sera elicited significant protection against bacteremia (Fig. 3*A*), with cfu geometric means lower by 2.6 and 1.6 logs, respectively, than that of the control group. Only anti-D1 serum afforded significant protection against mortality, giving 100% survival rate, whereas santi-D4 gave a protective trend, with 60% survival rate. These results indicate that immunization with D1 domain and at a lower extent D4 domain, elicits functional antibodies that may play a role in the protection.

D1 and D4 Antisera Recognize the Native Pilus and Linear Epitopes within RrgB—To investigate the differences in protective efficacy exerted by the isolated RrgB domains with respect to the FL RrgB in the *in vivo* assays, mouse sera were tested for their capability of recognizing the native pilus and the recombinant FL RrgB. Sera were probed against entire TIGR4 bacteria by FACS analysis. Sera raised against D1, D4, and D1+D2+D3+D4 gave a fluorescence intensity comparable with that obtained with anti-FL RrgB, whereas D2- and D3-specific antisera recognized the native pilus less effectively. To explain the lower recognition efficacy observed with D2 and D3, the same sera were titrated in ELISA against the recombinant FL RrgB. As shown in supplemental Fig. S3, antibody titers elicited by D2 and D3 immunization toward FL RrgB were about 10 times lower than those obtained by immunization with D1, D4, D1+D2+D3+D4, and FL RrgB, consistent with the results obtained by FACS analysis on the polymerized native RrgB.

To gain more insights on the epitopes recognized by the protective D1 and D4 polyclonal antibodies, a PepScan approach, suitable for identifying linear epitopes, was applied. Overlapping 15-mer peptides with an offset of 5 residues were synthesized *in situ* on a glass fiber membrane. The library of peptides tested covered residues 25–190 of D1 and 444–628 of D4. Incubation with the anti-D1 serum (previously used in passive protection experiments) revealed a single linear epitope covering residues 40–59 (D1-1) (Fig. 4*B*), whereas anti-D4 serum

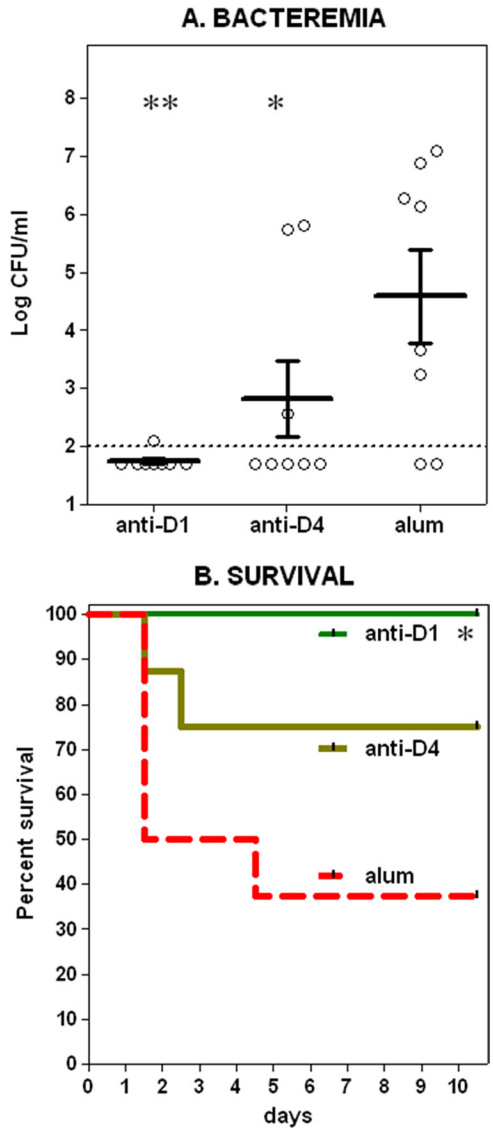


FIGURE 3. **Anti-RrgB D1 and D4 sera are protective in passive serum transfer experiments.** *Symbols* are described in the Fig. 2 legend. $n = 8$ for each of the groups.

detected a unique linear epitope (D4-1) spanning residues 494–508 (Fig. 4*C*).

Solution Structure of the D1 Domain—The solution structure of D1 was investigated by NMR spectroscopy. Its ^{1}H-^{15}N HSQC spectrum showed well dispersed resonances indicative of an overall well folded protein (supplemental Fig. S4). D1 showed a common IgG-like β sandwich fold (41 Å × 48 Å × 30 Å) and a topology of secondary structure elements drawn in Fig. 5. The core of the structure was formed by seven parallel and antiparallel β-strands: β_1(36–39), β_4(80–85), β_7(119–121), β_8(127–130), β_9(138–143), β_{10}(166–169), and β_{11}(178–180). These β-strands were arranged in two sheets (comprising β_1, β_8, β_{11} and β_4, β_7, β_9, β_{10}, respectively) packed against each other and flanked by two long segments (40–78, 87–117) located between strands β_1 and β_4 and strands β_4 and β_7, respectively. An α-helix (49–57), flanked by two short β-strands β_2 (42–44) and β_3 (73–75), was inserted within the first segment. Two additional β-strands β_5 (89–91) and β_6 (97–101) formed a β-hairpin structure inserted within the sec-

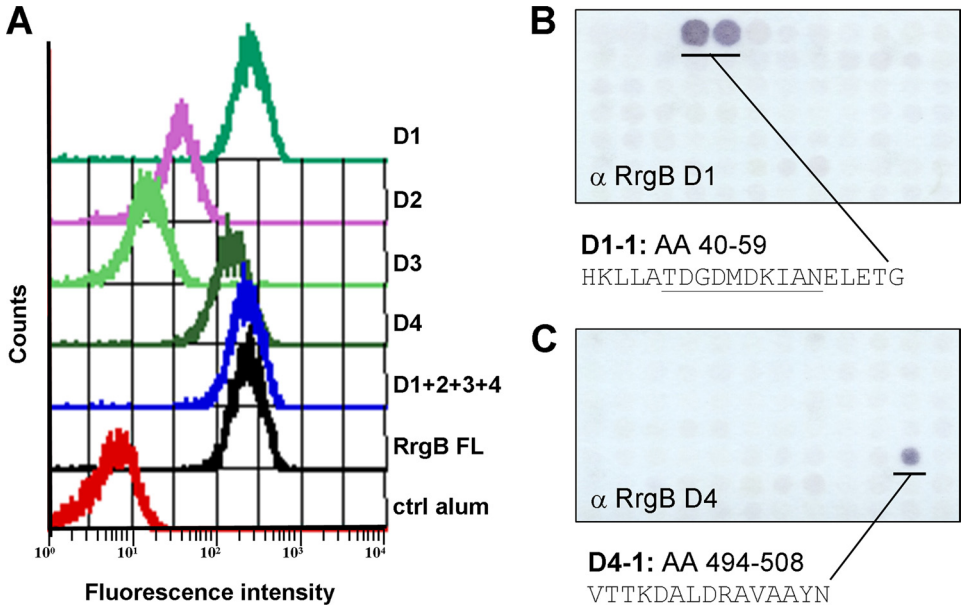


FIGURE 4. **Polyclonal antibodies raised against D1 and D4 recognize the native pilus and linear epitopes within RrgB efficiently.** *A*, TIGR4 bacteria were incubated with mouse primary antibodies directed against the specified recombinant proteins (1:400 dilution) followed by FITC-conjugated goat anti-mouse IgG (1:100 dilution). Bacterial staining was analyzed by flow cytometry (FACS-Calibur). Mouse control sera (immunized with PBS plus alum) were used as negative control. *B* and *C*, glass fiber membranes with arrayed peptides synthesized *in situ* covering residues 25–190 (D1) and 444–628 (D4) of RrgB were incubated with anti-D1 (*A*) or anti-D4 (*B*) polyclonal mouse antibodies (1:3000) and then with goat anti-mouse IgG alkaline phosphatase-conjugated antibodies (1:5000). Linear epitopes corresponding to peptide sequences recognized by the antibodies are reported. *Underlining* marks common residues present in adjacent peptides in the PepScan.

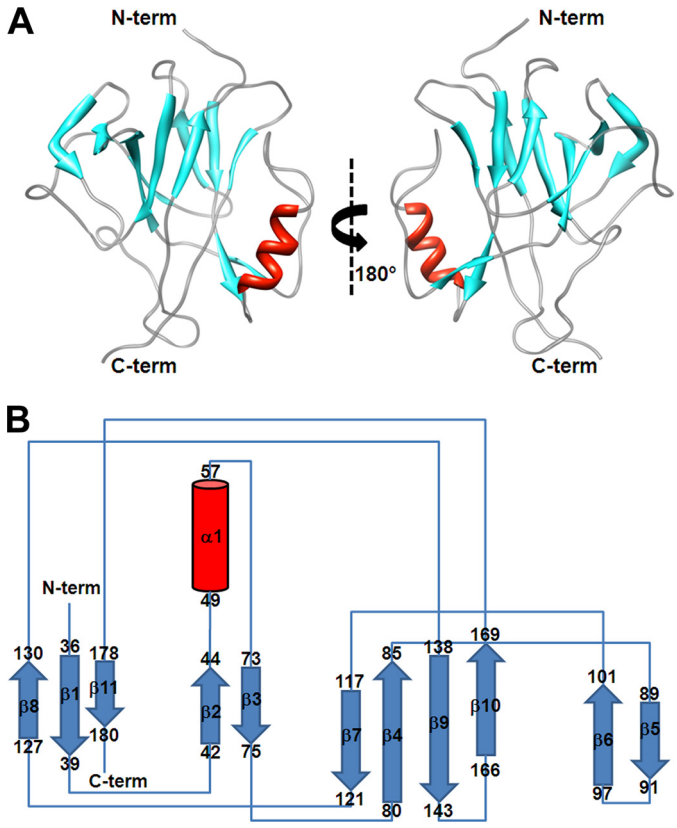


FIGURE 5. **Solution structure of the D1 domain.** *A*, ribbon diagram with the secondary structure elements. β-Strands are shown in *cyan*, the α-helix in *red*. *B*, topology diagram. The α-helix is represented by a *red cylinder*, and the β-strands are *cyan arrows*.

ond segment. In 50% of 20 conformers of the D1 family an additional β-sheet was formed by two short hydrogen-bonded β-strands (stretches 161–163 and 184–187).

The structure was well defined with the exception of three long loops corresponding to the stretches 56–69, 148–162, and 173–177 (supplemental Fig. S1). Heteronuclear relaxation measurements revealed that residues in the first two loops had heteronuclear NOE and longitudinal R_1 values lower and higher than average, respectively (supplemental Fig. S2). This behavior was a consequence of local internal motions occurring on a faster time scale with respect to the overall reorientational correlation time (τ_r) of the molecule and accordingly, a correlation time (τ_e) for these fast motions can be fitted for the previous mentioned loops (supplemental Fig. S2). In addition, conformational exchange processes, occurring on the millisecond-microsecond scale, affected some residues located in the region 155–164, as monitored by transverse R_2 relaxation rates higher than the average (supplemental Fig. S2). These data indicated that these loops experience higher flexibility, showing accordingly a low number of long ranges ^{1}H-^{1}H NOEs (supplemental Fig. S5). For a few residues, the assignment of the backbone resonances was also not achieved (Glu-143, His-145, Ser-146, Ser-148, Thr-149, Tyr-150, Val-152, and Gly-160), likely as a consequence of an increased local mobility. The loop between strands β_9 and β_{10} (residues 148–162) was the most disordered, with a r.m.s.d. value of about 2 Å (supplemental Fig. S1). Conformational exchange processes on the millisecond-microsecond time scale are observed also for the loop between strands β_{10} and β_{11} comprising residues 174–178.

The protein core is characterized by hydrophobic interactions between residues located on the first (strands β_1, β_8, β_{11}) and second (strands β_4, β_7, β_9, β_{10}) sheets. A salt bridge between two complementarily charged side chains of residues Lys-41 and Glu-143 was also present. The aliphatic side chains of residues Met-48, Ile-52, Ala-53, and Leu-56, all located on one side of the α-helix, formed hydrophobic interactions with

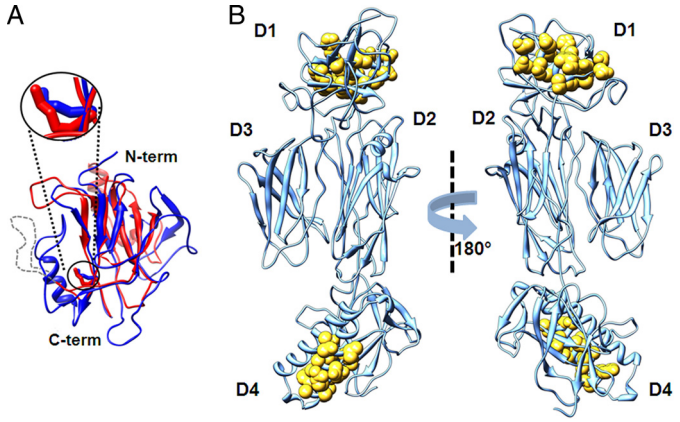


FIGURE 6. **Analysis of RrgB linker flexibility through superimposition and modeling.** *A*, superimposition of the D1 domain (*blue*) and the *C. diphtheriae* SpaA N-terminal domain (*red*). The position of Lys-190 residue, forming the intermolecular isopeptide bond between two neighboring SpaA molecules, is shown along with residue Lys-183 of RrgB, which occupies a similar position, and is indicated in a *black circle*. The missing loop in the SpaA crystal structure is indicated as a *gray dotted line*. *B*, modeled conformation of the full-length RrgB molecule obtained by combining the D1 NMR coordinates with the RrgB D2-D4 x-ray coordinates. D1-1 and D4-1 epitopes (residues 40–59 and 494–508 respectively) are rendered as *yellow spheres*.

aliphatic residues of β_2 and β_3 strands; these interactions determined the position of the helix with respect to the rest of the protein.

A search for related protein structures performed through the DALI Server (48) retrieved the N-terminal domain of the SpaA pilus backbone protein of *C. diphtheriae* (Protein Data Bank (PDB) code 3HTL; r.m.s.d. 2.1 Å), the C-terminal CNA3 domain of the major pilin protein of *Bacillus cereus* BcpA (PDB code 3KPT; r.m.s.d. 2.4 Å), and the N1 domain of the *Streptococcus agalactiae* minor pilin GBS52 (PDB code 2PZ4; r.m.s.d. 3.4 Å). Like D1, none of these domains contained intramolecular isopeptide bonds. However, only the SpaA domain is located at the N terminus of pilus backbone protein as D1; their overlay is presented in Fig. 6*A*.

In an attempt to determine the orientation of D1 with respect to the D2–D4 RrgB portion, a rigid body fitting of all the domains into the shape of the native pilus obtained from cryo-electron microscopy (cryo-EM) was attempted (24). Initial rigid body fitting of the D2–D4 crystal structure into the cryo-EM density map of the native pilus left an apical unoccupied volume, likely due to the absence of D1. However, when the simultaneous fitting of the D1 and the D2–D4 domains was carried out, some portions of D1 could not be accommodated into the apical empty volume (supplemental Fig. S6 and supplemental Table S6). To model the FL RrgB, we merged the D1 and D2–D4 structures into a single molecule by overlapping residues 188–191, shared by the C terminus of D1 and the N terminus of D2–D4. The relative orientation of D1 with respect to the others domains was then varied to best fit into the cryo-EM map (Fig. 6*B*). The final model, which represents only one of the possible orientations of D1 with respect to the rest of the protein, does not present steric clashes between D1 and the D2–D4 domains. The two linear epitopes, previously identified to be located within the D1 and D4 domains by peptide hybridization with specific protective polyclonal antibodies (Fig. 4), when mapped onto the FL model of RrgB, confirmed their superficial localization (Fig. 6*B*).

Lysine 183 of D1 Is Required for Intermolecular Isopeptide Bond Formation and Pilus Polymerization—Pili of Gram-positive bacteria are polymerized by means of intermolecular isopeptide bonds occurring between the Thr of the C-terminal LP*X*TG motif of a RrgB molecule and a Lys located at the N terminus of the following molecule (18, 29, 30). To identify the Lys implicated in the intermolecular isopeptide bond formation between two consecutive RrgB monomers, a TIGR4 RrgB deletion mutant (no longer able to assemble a pilus on its surface) was created. Subsequently, RrgB expression and pilus polymerization were restored in TIGR4ΔRrgB by transforming the mutant with plasmids expressing either wild-type RrgB or RrgB mutated in single Lys residues (Lys → Ala substitutions). Sequence analysis revealed that the D1 contains the canonical ^{181}YPKN^{184} pilin motif, with Lys-183 nicely superimposing onto the functional SpaA Lys-190, as shown in Fig. 6*A*. Noteworthy, D1 also presents the sequence fragment ^{160}GSKAVP^{165}, similar to the motif containing the Lys residue functional for the pilus assembly in *S. pyogenes* Spy0128 (31). Lys-183 and Lys-162, along with two additional Lys residues (Lys-138 and Lys-309), located on D1 and D3 domains, respectively, were selected and mutated.

WB analysis performed on whole cell bacterial lysates using rabbit polyclonal antisera raised against RrgA, RrgB, or RrgC revealed that all of the pilus proteins were expressed in all of the complemented mutants (Fig. 5). However, the typical pilus-associated high molecular weight ladder was revealed only in the case of TIGR4ΔRrgB complemented with RrgB wild type or RrgB mutated in Lys-162, Lys-138, or Lys-309. In contrast, in the mutant complemented with K183A, RrgB was detectable only as a monomer (Fig. 7), clearly indicating that Lys-183 is the residue implicated in intermolecular isopeptide bond formation.

Moreover, in the presence of the RrgB K183A substitution only hetero-oligomers composed of RrgA-RrgC (about 160 kDa) or RrgB-RrgC (about 110 kDa) could be detected by WB analysis (see *arrows* in Fig. 7, *A* and *C*). The absence of RrgA-RrgB heterodimers suggests that RrgA and RrgB are linked exclusively through intermolecular isopeptide bonds involving the D1 residue Lys-183 of RrgB.

DISCUSSION

Ever since their initial discovery, pili of Gram-positive bacteria have received considerable attention because they are associated with a number of different virulence mechanisms (15, 16, 33) and elicit protection in animal models (32, 49–52). In particular, the *S. pneumoniae* pilus was found to be implicated in the initial attachment to epithelial cells (15, 16, 33), and the pilus components, being protective in mouse models of infection, are regarded as potential vaccine candidates (32, 33). The major pilin RrgB is organized into four Ig-like domains (D1–D4), as shown by combined structural approaches (Ref. 24 and this work). We have tested the four RrgB domains in animal model experiments and have shown that the protective efficacy exerted by the combination of the four RrgB domains D1+D2+D3+D4 is comparable with that afforded by the FL

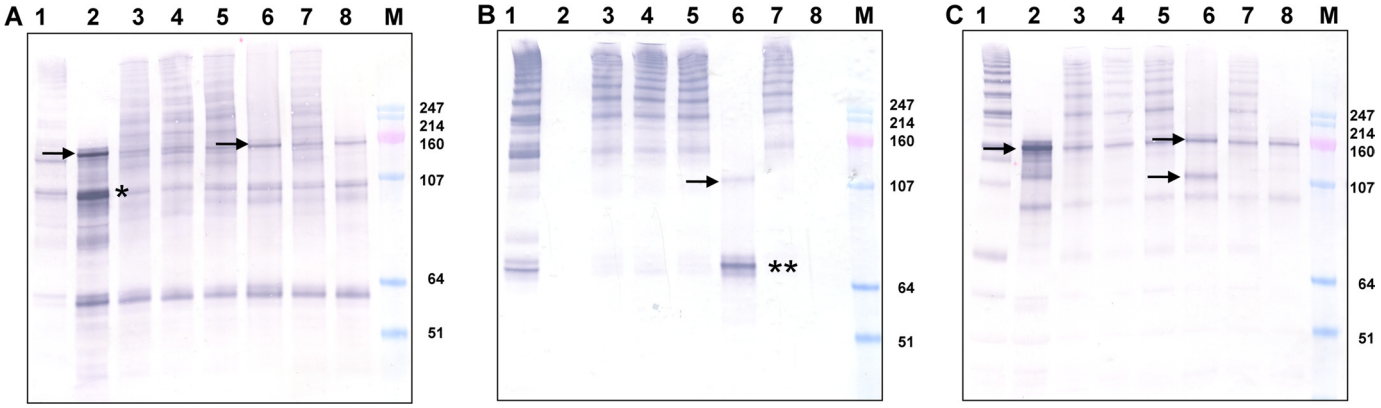


FIGURE 7. **Lys-183 of RrgB is implicated in intermolecular isopeptide bond formation.** WB analysis was performed using polyclonal rabbit antisera against RrgA (*A*), RrgB (*B*), and RrgC (*C*). In all *panels lanes* are loaded as follows: *1*, T4 WT; *2*, T4ΔRrgB; *3*, T4ΔRrgB▼RrgBWT; *4*, T4ΔRrgB▼RrgB(K138A); *5*, T4ΔRrgB▼RrgB(K162A); *6*, T4ΔRrgB▼RrgB(K183A); *7*, T4ΔRrgB▼RrgB(K309A); *8*, T4ΔRrgB▼pMU1328; *9*, molecular mass marker. *, RrgA monomer (estimated molecular mass of native monomeric RrgA is 92 kDa). **, RrgB monomer (estimated molecular mass of native monomeric RrgB is 65 kDa). *Arrows*, the band migrating at an apparent molecular mass of 160 kDa indicated by an *arrow* in *lane 2* and *6* of *A*, and *C* is compatible with an RrgA+RrgC complex; the band at 110 kDa present in *lane 6* of *B* and *C* could correspond to an RrgC+RrgB complex.

RrgB. Among the single domains, D1 is the most protective, and D4 retained an important part of the protective efficacy of the FL protein. The lower protection achieved by D2 and D3 compared with the FL protein, as well as with the D1 and D4, is probably because the antibodies elicited by the former two domains recognize the FL protein less efficiently, in both its native and recombinant forms.

This may be the result of smaller exposed surface areas experienced by D2 and D3 in the FL RrgB compared with the D2 and D3 isolated domains and with D1 and D4. It is possible, in fact, that the antibodies generated against the isolated domains are recognizing areas of D2 and D3 that are buried in the FL protein. Alternatively, D2 and D3 could assume a slightly different conformation when expressed as single domains, thus generating nonfunctional antibodies. On the other hand, the two linear epitopes (Fig. 6*B*) identified within D1 and D4 by PepScan analysis performed with protective polyclonal antibodies raised against the two domains (conformational epitopes are not detectable with this method) are well exposed on the surface of the RrgB molecule and could contribute to the protective activity exerted by D1 and D4. Further experiments are needed to understand to what extent these linear epitopes contribute to the protective activity exerted by the two domains. Taken together, these results suggest that RrgB contains multiple protective epitopes, thus confirming the potential of this vaccine candidate. Furthermore, although the existence of possible conformational epitopes involving residues from different domains cannot be excluded, their contribution to the overall protective efficacy might not be essential.

To obtain more information about the structural role of D1 and to try to correlate it with the protection data discussed above, we solved the solution structure of this domain by using NMR spectroscopy. D1 shows an Ig-like fold, does not contain any intramolecular isopeptide bond, and has many flexible regions. The observed D1 flexibility, indeed, could play a fundamental role in the specific antigen-antibody recognition process (53), thus accounting for D1 enhanced protection capability with respect to the more rigid D2–D4 domains, each one containing an intramolecular isopeptide bond. In fact, the protein structural plasticity could be related to the ability of D1 to undergo local conformational changes and to adapt its structure to optimize the interactions with the antibodies and increase the affinity and the specificity of the antigen-antibody recognition process. The dynamics of D1 could therefore strongly contribute to the interface adaptation for molecular recognition such that the antibody can select an optimal conformer from a wide distribution of possible D1 conformations. The rigid structure of the D2–D4 region prevents such an effective conformational selection for these domains. The above described phenomenon has been observed for other protein-protein or protein-DNA interaction processes (53, 54).

To shed light on the molecular mechanism driving pilus polymerization in *S. pneumoniae*, we investigated which of the lysine residues of D1 was engaged in the intermolecular isopeptide bond formation. Site-directed mutagenesis followed by complementation identified Lys-183 as crucial for the pilus assembly. This result is consistent with the observation that the spatial position of RrgB Lys-183 can be superimposed onto Lys-190 of the *C. diphtheriae* pilus backbone subunit SpaA, known to be involved in the intermolecular isopeptide linkage (30). Interestingly, as shown in Fig. 6*A*, both lysines are not fully available to form an external bond. In particular, the average relative solvent accessibility of the Lys-183 over the D1 family of conformers is 27.2% ± 4.8, as the Lys side chain projects into a cleft between the main body of the protein and the segment 40–78 containing the mobile loop 56–69 (27). This suggests that pilus backbone proteins, to be polymerized, might undergo conformational changes, probably involving not only the Lys residue (Lys-183) but also the flexible regions spatially close to it (49–69, 152–167, and 183–193), to allow the formation of the covalent intermolecular isopeptide bond. NMR mobility data indicate that the C terminus of the D1, where Lys-183 is located, and other loops are highly mobile and that such dynamics could be relevant for the intermolecular isopeptide bond formation (supplemental Fig. S2). Consistently, the absence of stabilizing intramolecular isopeptide bonds renders D1, unlike the other domains, less rigid and prone to conformational rearrangements. Furthermore, the occurrence of a struc-

tural rearrangement of D1 within the native pilus structure is also in line with the partial fitting of its NMR structure onto the molecular shape of the native pilus determined by cryo-EM.

Mutagenesis and complementation data were used to analyze the covalent links established among the three pilus proteins either in the absence of RrgB or in the presence of the RrgB K183A mutant, in an attempt to provide further insights into the possible organization of the native pilus. In the presence of the nonpolymerizing RrgB K183A mutant, the lack of RrgA-RrgB heterodimers provides evidence that RrgA and RrgB are unidirectionally linked only through the Lys-183 of RrgB and the C-terminal Thr of RrgA. This arrangement is in accordance with the model proposed by Hilleringmann *et al.*, positioning RrgA at the pilus terminus, thus ruling out the alternative possibility of RrgA being incorporated along the pilus shaft (22). Conversely, an RrgA-RrgC multimer is detectable either when RrgB is not expressed, as already reported by Falker *et al.* (20) and Le Mieux *et al.* (55), or in the presence of nonpolymerizing RrgB K183A. In this case, the mutated RrgB is competing with RrgA for the linkage to RrgC, as demonstrated by the presence of RrgB-RrgC hetero-oligomers (Fig. 7, *B* and *C*). Concomitant detection of RrgA-RrgC hetero-oligomers even under these conditions further strengthens the idea that in wild-type bacteria, although not detectable by electron microscopy analysis of purified pili, a fraction of RrgA and RrgC might be directly linked to each other.

In conclusion, this study provides additional information elucidating pilus proteins features and also paves the way to the rational design of new RrgB-based molecules to implement a protein-based vaccine against pneumococcal disease. Moreover, the newly acquired structural and dynamic information on the RrgB molecule provided by this study suggests that the conformational flexibility of D1 is pivotal for the protein-antibody recognition process. These findings together with the new functional information could be used to better understand pilus functions and its role in pathogenesis.

Acknowledgments—We thank Giacomo Matteucci and Tommaso Pasquali, who managed Animal Resources; Marco Tortoli, Stefania Torricelli, Luigi Manganelli, and Elena Amantini, who took care of the animal treatments; Silvia Maccari, Esmeralda Bizzarri, and Alessia Corrado, who lent technical assistance for the in vivo experiments; and Morena Lo Sapio for technical assistance in FACS and ELISA experiments. Intercell AG kindly provided us with pMU1328. Luisa Lozzi and Luisa Bracci (University of Siena) synthesized the peptides onto the glass fiber membranes used for Pepscan analysis. Finally, we thank Mariagrazia Pizza for a critical review of the manuscript.

REFERENCES

1. Fletcher, M. A., and Fritzell, B. (2007) *Vaccine* **25,** 2507–2512
2. Kim, K. S. (2010) *Lancet Infect. Dis.* **10,** 32–42
3. O'Brien, K. L., Wolfson, L. J., Watt, J. P., Henkle, E., Deloria-Knoll, M., McCall, N., Lee, E., Mulholland, K., Levine, O. S., and Cherian, T. (2009) *Lancet* **374,** 893–902
4. Pelton, S. I., and Leibovitz, E. (2009) *Pediatr. Infect. Dis. J.* **28,** S133–137
5. Ryan, M. W., and Antonelli, P. J. (2000) *Laryngoscope* **110,** 961–964
6. van der Poll, T., and Opal, S. M. (2009) *Lancet* **374,** 1543–1556
7. Melegaro, A., Gay, N. J., and Medley, G. F. (2004) *Epidemiol. Infect.* **132,** 433–441
8. Melegaro, A., Choi, Y. H., George, R., Edmunds, W. J., Miller, E., and Gay, N. J. (2010) *BMC Infect. Dis.* **10,** 90
9. Millar, E. V., O'Brien, K. L., Zell, E. R., Bronsdon, M. A., Reid, R., and Santosham, M. (2009) *Pediatr. Infect. Dis. J.* **28,** 711–716
10. Barocchi, M. A., Ries, J., Zogaj, X., Hemsley, C., Albiger, B., Kanth, A., Dahlberg, S., Fernebro, J., Moschioni, M., Masignani, V., Hultenby, K., Taddei, A. R., Beiter, K., Wartha, F., von Euler, A., Covacci, A., Holden, D. W., Normark, S., Rappuoli, R., and Henriques-Normark, B. (2006) *Proc. Natl. Acad. Sci. U.S.A.* **103,** 2857–2862
11. Dramsi, S., Caliot, E., Bonne, I., Guadagnini, S., Prévost, M. C., Kojadinovic, M., Lalioui, L., Poyart, C., and Trieu-Cuot, P. (2006) *Mol. Microbiol.* **60,** 1401–1413
12. Lauer, P., Rinaudo, C. D., Soriani, M., Margarit, I., Maione, D., Rosini, R., Taddei, A. R., Mora, M., Rappuoli, R., Grandi, G., and Telford, J. L. (2005) *Science* **309,** 105
13. LeMieux, J., Hava, D. L., Basset, A., and Camilli, A. (2006) *Infect. Immun.* **74,** 2453–2456
14. Mora, M., Bensi, G., Capo, S., Falugi, F., Zingaretti, C., Manetti, A. G., Maggi, T., Taddei, A. R., Grandi, G., and Telford, J. L. (2005) *Proc. Natl. Acad. Sci. U.S.A.* **102,** 15641–15646
15. Hilleringmann, M., Giusti, F., Baudner, B. C., Masignani, V., Covacci, A., Rappuoli, R., Barocchi, M. A., and Ferlenghi, I. (2008) *PLoS Pathog.* **4,** e1000026
16. Nelson, A. L., Ries, J., Bagnoli, F., Dahlberg, S., Fälker, S., Rounioja, S., Tschöp, J., Morfeldt, E., Ferlenghi, I., Hilleringmann, M., Holden, D. W., Rappuoli, R., Normark, S., Barocchi, M. A., and Henriques-Normark, B. (2007) *Mol. Microbiol.* **66,** 329–340
17. Telford, J. L., Barocchi, M. A., Margarit, I., Rappuoli, R., and Grandi, G. (2006) *Nat. Rev. Microbiol.* **4,** 509–519
18. Ton-That, H., and Schneewind, O. (2004) *Trends Microbiol.* **12,** 228–234
19. El Mortaji, L., Terrasse, R., Dessen, A., Vernet, T., and Di Guilmi, A. M. (2010) *J. Biol. Chem.* **285,** 12405–12415
20. Fälker, S., Nelson, A. L., Morfeldt, E., Jonas, K., Hultenby, K., Ries, J., Melefors, O., Normark, S., and Henriques-Normark, B. (2008) *Mol. Microbiol.* **70,** 595–607
21. Manzano, C., Contreras-Martel, C., El Mortaji, L., Izoré, T., Fenel, D., Vernet, T., Schoehn, G., Di Guilmi, A. M., and Dessen, A. (2008) *Structure* **16,** 1838–1848
22. Hilleringmann, M., Ringler, P., Müller, S. A., De Angelis, G., Rappuoli, R., Ferlenghi, I., and Engel, A. (2009) *EMBO J.* **28,** 3921–3930
23. Izoré, T., Contreras-Martel, C., El Mortaji, L., Manzano, C., Terrasse, R., Vernet, T., Di Guilmi, A. M., and Dessen, A. (2010) *Structure* **18,** 106–115
24. Spraggon, G., Koesema, E., Scarselli, M., Malito, E., Biagini, M., Norais, N., Emolo, C., Barocchi, M. A., Giusti, F., Hilleringmann, M., Rappuoli, R., Lesley, S., Covacci, A., Masignani, V., and Ferlenghi, I. (2010) *PLoS One* **5,** e10919
25. Budzik, J. M., Poor, C. B., Faull, K. F., Whitelegge, J. P., He, C., and Schneewind, O. (2009) *Proc. Natl. Acad. Sci. U.S.A.* **106,** 19992–19997
26. Kang, H. J., and Baker, E. N. (2009) *J. Biol. Chem.* **284,** 20729–20737
27. Kang, H. J., Paterson, N. G., Gaspar, A. H., Ton-That, H., and Baker, E. N. (2009) *Proc. Natl. Acad. Sci. U.S.A.* **106,** 16967–16971
28. Alegre-Cebollada, J., Badilla, C. L., and Fernández, J. M. (2010) *J. Biol. Chem.* **285,** 11235–11242
29. Guttilla, I. K., Gaspar, A. H., Swierczynski, A., Swaminathan, A., Dwivedi, P., Das, A., and Ton-That, H. (2009) *J. Bacteriol.* **191,** 5603–5612
30. Ton-That, H., and Schneewind, O. (2003) *Mol. Microbiol.* **50,** 1429–1438
31. Kang, H. J., Coulibaly, F., Clow, F., Proft, T., and Baker, E. N. (2007) *Science* **318,** 1625–1628
32. Gianfaldoni, C., Censini, S., Hilleringmann, M., Moschioni, M., Facciotti, C., Pansegrau, W., Masignani, V., Covacci, A., Rappuoli, R., Barocchi, M. A., and Ruggiero, P. (2007) *Infect. Immun.* **75,** 1059–1062
33. Moschioni, M., Emolo, C., Biagini, M., Maccari, S., Pansegrau, W., Donati, C., Hilleringmann, M., Ferlenghi, I., Ruggiero, P., Sinisi, A., Pizza, M., Norais, N., Barocchi, M. A., and Masignani, V. (2010) *Infect. Immun.* **78,** 5033–5042
34. Achen, M. G., Davidson, B. E., and Hillier, A. J. (1986) *Gene* **45,** 45–49
35. Alloing, G., Martin, B., Granadel, C., and Claverys, J. P. (1998) *Mol. Microbiol.* **29,** 75–83

36. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR* **13,** 289–302
37. Herrmann, T., Güntert, P., and Wüthrich, K. (2002) *J. Mol. Biol.* **319,** 209–227
38. Pearlman, A. D., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatman, T. E., III, DeBolt, S., Ferguson, D., Seibel, G., and Kollman, P. (1995) *Comp. Phys. Commun.* **91,** 1–41
39. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graph.* **14,** 29–32
40. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) *Biochemistry* **33,** 5984–6003
41. Grzesiek, S., and Bax, A. (1993) *J. Am. Chem. Soc.* **115,** 12593–12594
42. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc.* **104,** 4546–4559
43. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc.* **104,** 4559–4570
44. Dosset, P., Hus, J. C., Blackledge, M., and Marion, D. (2000) *J. Biomol. NMR* **16,** 23–28
45. Goddard, T. D., Huang, C. C., and Ferrin, T. E. (2007) *J. Struct. Biol.* **157,** 281–287
46. Meng, E. C., Pettersen, E. F., Couch, G. S., Huang, C. C., and Ferrin, T. E. (2006) *BMC Bioinformatics* **7,** 339
47. Goddard, T. D., Huang, C. C., and Ferrin, T. E. (2005) *Structure* **13,** 473–482
48. Holm, L., and Sander, C. (1996) *Methods Enzymol.* **266,** 653–662
49. Abbot, E. L., Smith, W. D., Siou, G. P., Chiriboga, C., Smith, R. J., Wilson, J. A., Hirst, B. H., and Kehoe, M. A. (2007) *Cell. Microbiol.* **9,** 1822–1833
50. Pezzicoli, A., Santi, I., Lauer, P., Rosini, R., Rinaudo, D., Grandi, G., Telford, J. L., and Soriani, M. (2008) *J. Infect. Dis.* **198,** 890–898
51. Rosini, R., Rinaudo, C. D., Soriani, M., Lauer, P., Mora, M., Maione, D., Taddei, A., Santi, I., Ghezzo, C., Brettoni, C., Buccato, S., Margarit, I., Grandi, G., and Telford, J. L. (2006) *Mol. Microbiol.* **61,** 126–141
52. Margarit, I., Rinaudo, C. D., Galeotti, C. L., Maione, D., Ghezzo, C., Buttazzoni, E., Rosini, R., Runci, Y., Mora, M., Buccato, S., Pagani, M., Tresoldi, E., Berardi, A., Creti, R., Baker, C. J., Telford, J. L., and Grandi, G. (2009) *J. Infect. Dis.* **199,** 108–115
53. Mittag, T., Kay, L. E., and Forman-Kay, J. D. (2010) *J. Mol. Recognit.* **23,** 105–116
54. Dyson, H. J., and Wright, P. E. (2005) *Nat. Rev. Mol. Cell Biol.* **6,** 197–208
55. LeMieux, J., Woody, S., and Camilli, A. (2008) *J. Bacteriol.* **190,** 6002–6013

Prevalence of pilus-encoding islets among acute otitis media *Streptococcus pneumoniae* isolates from Israel

M. Moschioni[1], G. De Angelis[1], S. Melchiorre[1], V. Masignani[1], E. Leibovitz[2], M. A. Barocchi[1] and R. Dagan[2,3]

1) *Novartis Vaccines and Diagnostics s.r.l., Research Center, Siena, Italy,* 2) *Faculty of Health Sciences, Ben-Gurion University of the Negev and* 3) *Pediatric Infectious Disease Unit, Soroka University Medical Center, Beer-Sheva, Israel*

Abstract

We evaluated the distribution of the two known *Streptococcus pneumoniae* pilus encoding islets (PI-1 and PI-2) among a panel of 113 acute otitis media clinical isolates from Israel. PI-1 was present in 30.1% (n = 34) of the isolates tested, and PI-2 was present in 7% (n = 8). In addition, we found that: (i) the PI positive isolates, 50% of which belong to the international clones Spain^{9V}-3 (ST156) and Taiwan^{19F}-14 (ST236), correlate with the genotype (as determined by multilocus sequence typing) but not with the serotype; (ii) PI-2 was not present in the absence of PI-1; and (iii) the frequency of PI-1 was higher among antibiotic-resistant isolates.

Keywords: Acute otitis media, antibiotic resistance, genotype, pilus-encoding islet, *Streptococcus pneumoniae*

Original Submission: 19 August 2009; **Revised Submission:** 22 October 2009; **Accepted:** 23 October 2009
Editor: J.-L. Mainardi
Article published online: 2 November 2009

Clin Microbiol Infect 2010; **16:** 1501–1504
10.1111/j.1469-0691.2010.03105.x

Corresponding author: R. Dagan, Faculty of Health Sciences, Ben-Gurion University of the Negev and Pediatric Infectious Disease Unit, Soroka University Medical Center, Beer-Sheva, Israel
E-mail: rdagan@bgu.ac.il

The human pathogen *Streptococcus pneumoniae* is commonly associated with invasive diseases such as meningitis and sepsis. In addition, pneumococci are the most frequent cause of upper respiratory infections such as sinusitis and acute otitis media (AOM), which is one of the most widespread childhood infections and a major cause of morbidity in children [1,2]. Hence, it is important to investigate the presence of pathogenic factors that may be responsible for disease outcome among AOM clinical isolates.

Recently, genetic analysis of *S. pneumoniae* clinical isolates demonstrated that they harbour pilus structures encoded by the *rlrA* pathogenicity islet (pilus islet-1, PI-1) and the pilus islet-2 (PI-2). The PI-1 pilus was shown to be involved in virulence [3,4] and antibodies raised against its protein subunits were protective in a murine model of infection [5]. However, the islets are not widely distributed in *S. pneumoniae*. Three independent studies demonstrated that, in both invasive and nasopharyngeal clinical isolates, the frequency of the PI-1 islet is approximately 30% [6–8], although the incidence was higher among antibiotic-resistant clones [8]. This suggests that the interplay between the pilus (encoded by PI-1) and antibiotic resistance may facilitate the global spread of antibiotic nonsusceptible pneumococci [9]. A recent study demonstrated that PI-2 is present in approximately 16% of invasive and nasopharyngeal clinical isolates [10], and that both islets are present in the Taiwan^{19F}-14 (ST236) clone, whose spread is responsible for the increasing incidence of antibiotic-resistant isolates in many countries [11–13].

Therefore, to evaluate a possible correlation between AOM and the presence of PIs, we aimed to assess the prevalence of the pilus encoding islets in a collection of AOM clinical isolates, for which no data are available thus far.

A total of 113 pneumococcal isolates were obtained from the middle ear fluid of 113 patients with AOM presenting at the Soroka University Medical Center Pediatric Emergency Department, from 1 January 2007 to 29 March 2007. The patients were Jewish children (n = 55) and Bedouin children (n = 58) who did not receive the heptavalent conjugate pneumococcal vaccine (PCV7) because PCV 7 vaccination was not implemented in Israel at the time of the study. The age of the children was in the range 0.8–96.9 months (mean ± SD 13.3 ± 15.5 months; median 9.7 months). Of the 113 isolates, 85 (75%) were obtained by tympanocentesis and 17 (15%) were from spontaneous drainage; it was not clear whether the remaining 11 isolates (10%) were obtained from tympanocentesis or spontaneous drainage. Of 107 isolates from which information about antibiotic treatment of the patient in the month prior to culture was available, 32 (30%) were from children receiving antibiotics in previous months. *S. pneumoniae* was isolated as a single pathogen in 62/113 (65%) cases; in 46 (41%) together with *Haemophilus*

influenzae; and in five (4%) together with other combinations of pathogens. Identification, serotyping and antimicrobial susceptibility testing was performed as described previously [14].

The most commonly isolated serotype was serotype 14, followed by 19F, 23F, 19A, 6B and 18C (Fig. 1a). All isolates were tested for the presence of PI-1 and PI-2 by PCR using a specific set of primers as described previously [8,10]. Amplifications to determine the genomic location and the presence of islets were performed directly from bacterial colonies grown overnight on blood-agar plates.

Of 113 isolates, 34 (30.1%) were positive for *S. pneumoniae* PI-1 and eight (7%) were positive for PI-2; PI-1 was found in serotypes 6A, 6B, 14, 19F, 23F, 33A and 11A, (Fig. 1a). On the other hand, PI-2 was only found in serotype 19F isolates and always associated with PI-1 (Fig. 1a). Subsequently, the isolates containing PI-1 or both PI-1 and PI-2 were typed by multilocus sequence typing [15] and grouped by E-BURST analysis (http://eburst.mlst.net/) by clonal complex (CC) (Fig. 1b), revealing that they belonged to CCs already reported to be most likely to carry the PI-1 [6–8]. Isolates that harboured both pilus islets were mainly of the CC271, single or double locus variants of the international clone Taiwan^{19F}-14. Among the isolates containing only PI-1, CC156 was the prevalent clone (international clone Spain^{9V}-3). Interestingly, out of 12 CC156 isolates, 11 (92%) were serotype 14, confirming the increasing association of this serotype with clone CC156 [9].

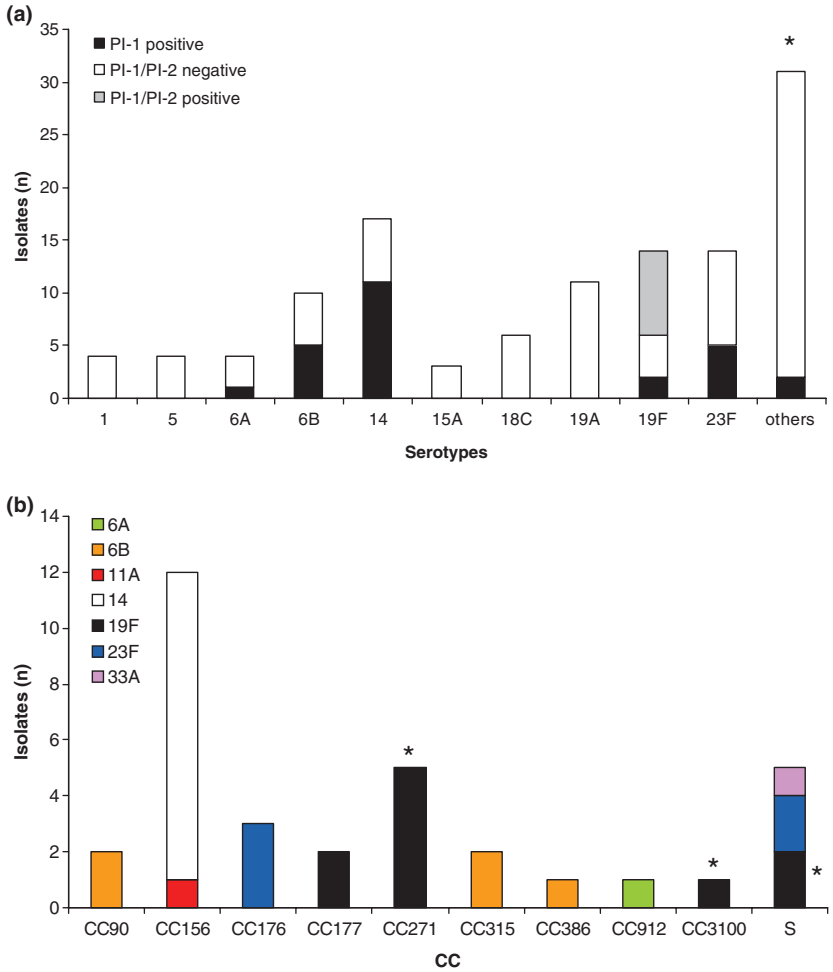


FIG. 1. Acute otitis media (AOM) isolates collection characterized for the presence of pilus islet (PI)-1 and PI-2. (a) Collection of AOM isolates stratified by serotype and analysed for the presence of PI-1 and PI-2. *Serotypes with two isolates only (3, 4, 8, 15B, 16F, 18A, 33A) or one isolate only (7B, 9N, 9V, 11A, 23A, 23B, 24F, 31, 33F, 35B, not typable). PI-1 was present only in one of the two serotype 33A isolates and in the 11A isolate. (b) PI-1 and PI-2 positive isolates stratified by clonal complex (CC), determined by E-BURST grouping, after multilocus sequence typing. The serotypes to which the CCs are associated are shown with different colours. CCs have been named according to the sequence typing number of the E-BURST predicted founder. S, singletons. *Serotype 19F isolates positive for both PI-1 and PI-2.

TABLE 1. Antibiotic resistance pattern in selected serotypes, in isolates with and without the pilus islet (PI)-1

	14		19F		23F		33A		6A		6B		Others		Total	
	PII (−)	PII (+)	PII (−)	PII (+)	PII (−)	PII (+)	PII (−)	PII (+)	PII (−)	PII (+)	PII (−)	PII (+)	PII (−)	PII (+)	PII(+)/resistant isolates	
	(*n* = 6)	(*n* = 11)	(*n* = 4)	(*n* = 10)	(*n* = 9)	(*n* = 5)	(*n* = 1)	(*n* = 1)	(*n* = 3)	(*n* = 1)	(*n* = 5)	(*n* = 5)	(*n* = 51)	(*n* = 1)		
Erythromycin resistance	5	2	1	8	0	0	1	0	2	1	4	5	2	0	16/31 (51.6%)	
Penicillin MIC ≥0.1 mg/L	5	11	1	9	6	4	0	0	2	0	3	2	21	1	27/65 (41.5%)	
Penicillin MIC ≥1.0 mg/L	3	2	1	4	4	0	0	0	0	0	0	0	2	0	6/16 (37.5%)	
Multidrug resistance*	5	1	1	8	2	0	0	0	0	1	3	4	2	0	14/27 (51.8%)	
Susceptible to all antibiotics	0	0	3	1	2	0	0	1	0	0	1	0	27	0	PII(+)/susceptible isolates 2/35(5.7%)	

*Resistant to ≥3 drug classes

No correlation was found with age or gender of the children, whereas the percentage of PI positive isolates among the Jewish population was higher (43.6%) compared with to Bedouins (17.2%), although the proportion of antibiotic resistant clones was comparable in the two groups. Moreover, children who had experienced one or more otitis media episodes in the previous year were more likely to be infected with *S. pneumoniae* carrying the pilus-encoding islets than children who had not (46% vs. 18%; p 0.002). The concomitant presence of *H. influenzae* in the infection site did not correlate with PIs positive *S. pneumoniae* (28%).

Because PI-1 was found mainly in serotypes with prevalent antibiotic resistance, we analysed its presence or absence by the antibiotic-resistance pattern in each of the serotypes in which it was described (Table 1). The percentage of PI-1 positive isolates among antibiotic-resistant strains reached 41.5% in penicillin-nonsusceptible isolates (MIC >0.1 mg/L), 51.8% among multidrug-resistant isolates, and 51.6% in erythromycin- resistant isolates. By contrast, the proportion of PI-1 strains in strains susceptible to all tested antibiotics was 5.7% (p 0.001). These findings confirm the previously found association between piliated *S. pneumoniae* and antibiotic resistance [8].

In conclusion, the data obtained in the present study indicate that the incidence of piliated pneumococci in otitis media isolates is not greater than in other invasive disease isolates and that the pilus does not appear to confer an additional selective benefit in the middle ear during otitis media. A limitation of the present study is the small sample size of clinical isolates tested. However, the results obtained confirm the correlation of PI in *S. pneumoniae* with genotype and antibiotic resistance. Further studies are needed to understand the epidemiological relevance of the pilus and its role in disease outcome. Although vaccination with PCV7 reduced the frequency of persistent and recurrent AOM [16], multidrug-resistant nonvaccine serotypes such as serotype 19A are increasing, associated with CC156 [17]. This clone was a prevalent clone in the present study. Therefore, the elucidation of the pathogenesis of AOM may help advance the development of a serotype-independent pneumococcal vaccine for otitis media, and a new generation protein vaccine that includes pilus subunits could reduce otitis media disease.

Transparency Declaration

M. Moschioni, G. De Angelis, S. Melchiorre, V. Masignani and M.A. Barocchi are employees of Novartis Vaccines and Diagnostics. E. Leibovitz has no conflict of interest. R. Dagan

has had the following financial interests and/or arrangements with the corporate organizations listed herebelow in the past 5 years: Grant/Research support – Berna/Crucell, Wyeth/Pfizer, MSD; Scientific Consultancy – Berna/Crucell, GlaxoSmithKline, Novartis, Wyeth/Pfizer, Protea, MSD; Speaker – Berna/Crucell, GlaxoSmithKline, Wyeth/Pfizer; Shareholder – Protea.

References

1. Hausdorff WP, Feikin DR, Klugman KP. Epidemiological differences among pneumococcal serotypes. *Lancet Infect Dis* 2005; 5: 83–93.
2. Rovers MM, Schilder AG, Zielhuis GA, Rosenfeld RM. Otitis media. *Lancet* 2004; 363: 465–473.
3. Barocchi MA, Ries J, Zogaj X *et al.* A pneumococcal pilus influences virulence and host inflammatory responses. *Proc Natl Acad Sci USA* 2006; 103: 2857–2862.
4. Nelson AL, Ries J, Bagnoli F *et al.* RrgA is a pilus-associated adhesin in *Streptococcus pneumoniae*. *Mol Microbiol* 2007; 66: 329–340.
5. Gianfaldoni C, Censini S, Hilleringmann M *et al. Streptococcus pneumoniae* pilus subunits protect mice against lethal challenge. *Infect Immun* 2007; 75: 1059–1062.
6. Aguiar SI, Serrano I, Pinto FR, Melo-Cristino J, Ramirez M. The presence of the pilus locus is a clonal property among pneumococcal invasive isolates. *BMC Microbiol* 2008; 8: 41.
7. Basset A, Trzcinski K, Hermos C *et al.* Association of the pneumococcal pilus with certain capsular serotypes but not with increased virulence. *J Clin Microbiol* 2007; 45: 1684–1689.
8. Moschioni M, Donati C, Muzzi A *et al. Streptococcus pneumoniae* contains 3 *rlr*A pilus variants that are clonally related. *J Infect Dis* 2008; 197: 888–896.
9. Sjostrom K, Blomberg C, Fernebro J *et al.* Clonal success of piliated penicillin nonsusceptible pneumococci. *Proc Natl Acad Sci USA* 2007; 104: 12907–12912.
10. Bagnoli F, Moschioni M, Donati C *et al.* A second pilus type in *Streptococcus pneumoniae* is prevalent in emerging serotypes and mediates adhesion to host cells. *J Bacteriol* 2008; 190: 5480–5492.
11. Hsieh YC, Chang KY, Huang YC *et al.* Clonal spread of highly beta-lactam-resistant *Streptococcus pneumoniae* isolates in Taiwan. *Antimicrob Agents Chemother* 2008; 52: 2266–2269.
12. Mavroidi A, Paraskakis I, Pangalis A *et al.* Spread of the *Streptococcus pneumoniae* Taiwan19F-14 clone among children in Greece. *Clin Microbiol Infect* 2007; 13: 1213–1216.
13. Moore MR, Gertz RE Jr, Woodbury RL *et al.* Population snapshot of emergent *Streptococcus pneumoniae* serotype 19A in the United States, 2005. *J Infect Dis* 2008; 197: 1016–1027.
14. Barkai G, Greenberg D, Givon-Lavi N, Dreifuss E, Vardy D, Dagan R. Community prescribing and resistant *Streptococcus pneumoniae*. *Emerg Infect Dis* 2005; 11: 829–837.
15. Enright MC, Spratt BG. A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. *Microbiology* 1998; 144: 3049–3060.
16. Casey JR, Pichichero ME. Changes in frequency and pathogens causing acute otitis media in 1995–2003. *Pediatr Infect Dis J* 2004; 23: 824–828.
17. Pichichero ME, Casey JR. Emergence of a multiresistant serotype 19A pneumococcal strain not included in the 7-valent conjugate vaccine as an otopathogen in children. *JAMA* 2007; 298: 1772–1778.

Modified sequential multiplex PCR for determining capsular serotypes of invasive pneumococci recovered from Seville

P. Iraurgui[1], M. J. Torres[2], A. Gandia[1], I. Vazquez[1], E. G. Cabrera[2], I. Obando[3], J. Garnacho[4] and J. Aznar[1,2]

1) *Microbiology Service, Hospital Universitario Virgen del Rocío,* 2) *Instituto de Biomedicina de Sevilla, Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla,* 3) *Paediatric Service and* 4) *Intensive Care Unit, Hospital Universitario Virgen del Rocío, Sevilla, Spain*

Abstract

The heptavalent pneumococcal vaccine's introduction resulted in a decline in invasive disease caused by *Streptococcus pneumoniae*, but was accompanied by an increase in non-vaccine serotypes. We evaluated a modified scheme of the sequential multiplex PCRs adapted to the prevalence of serotypes in Seville (Spain) for determining capsular serotypes of *S. pneumoniae* invasive clinical isolates. In adults, the modified scheme allowed us to type 73% with the first three reactions, and 92% with two additional PCRs. In paediatric patients, it allowed us to type 73.5% with the first three reactions, and 90% with the two additional PCRs. The multiplex PCR approach was successfully adapted to target the serotypes most prevalent in Seville.

Keywords: Capsular polysaccharide, invasive infection, multiplex PCR, serotype, *Streptococcus pneumoniae*

Original Submission: 22 July 2009; **Revised Submission:** 18 November 2009; **Accepted:** 19 November 2009
Editor: J.-L. Mainardi
Article published online: 7 December 2009

Clin Microbiol Infect 2010; **16:** 1504–1507
10.1111/j.1469-0691.2009.03129.x

Corresponding author: M. J. Torres, Departamento de Microbiología, Facultad de Medicina, Apdo. 914, 41080 Seville, Spain
E-mail: mjtorres@us.es

Streptococcus pneumoniae causes severe illnesses in the elderly and children. The immunochemistry of the capsular polysaccharide differentiates pneumococci into 91 distinct